RETURN TIMES, RECURRENCE DENSITIES AND ENTROPY FOR ACTIONS OF SOME DISCRETE AMENABLE GROUPS

MICHAEL HOCHMAN

Abstract. It is a theorem of Wyner & Ziv and Ornstein & Weiss that if one observes the initial k symbols $X_0,\ldots,X_{k-1}$ of a typical realization of a finite valued ergodic process with entropy h, the waiting time until this sequence appears again in the same realization grows asymptotically like 2^{hk} [7, 12]. A similar result for random fields was obtained in [8]: in this case one observes cubes in $\mathbb{Z}^d$ instead of initial segments.

In the present paper we describe generalizations of this. We examine what happens when the set of possible return times is restricted: Fix an increasing sequence of sets of possible times $\{W_n\}$ and define R_k to be the first n such that $X_0,\ldots,X_{k-1}$ recurs at some time in W_n. It turns out that $|W_{R_k}|$ cannot drop below 2^{hk} asymptotically. We obtain conditions on the sequence $\{W_n\}$ which ensure that $|W_{R_k}|$ is asymptotically equal to 2^{hk}.

We consider also recurrence densities of initial blocks and derive a uniform Shannon-McMillan-Breiman theorem: Informally, if $U_{k,n}$ is the density of recurrences of the block $X_0,\ldots,X_{k-1}$ in $X_{-n},\ldots,X_n$ then $U_{k,n}$ grows at a rate of 2^{hk}, uniformly in n. We examine the conditions under which this is true when the recurrence times are again restricted to some sequence of sets $\{W_n\}$

The above questions are examined in the general context of finite-valued processes parameterized by discrete amenable groups. We show that many classes of groups have time-sequences $\{W_n\}$ along which return times and recurrence densities behave as expected. An interesting feature here is that this can happen also when the time sequence lies in a small subgroup of the parameter group.

1. Introduction

1.1. **Background.** Let $\{X_n\}_{n\in\mathbb{Z}}$ be an ergodic process with values in some finite set Σ, defined on a probability space $(\Omega,\mathcal{F},P)$. Let h denote the entropy of the process. For a realization $x=(x_n)_{n\in\mathbb{Z}}$ of the process and for each k consider the finite word $x_0x_1\ldots x_{k-1}$ obtained by viewing x through the "window" consisting of coordinates $0,1,\ldots,k$. The first backward recurrence time of this word is defined as

$$R_k^-(x)=\min\{n\,:\,n>0\text{ and }x_{-n+i}=x_i\text{ for all }0\leq i<k\}$$

Set $R_k^-(x)=\infty$ if the sequence $x_0x_1\ldots x_{k-1}$ does not appear again in the past portion of x.

In [12], A. Wyner and J. Ziv proved that $\frac{1}{k}\log R_k^-\to h$ in probability and $\limsup_k\frac{1}{2k+1}\log R_k^-\leq h$ almost surely. This result was strengthened by D. Ornstein and B. Weiss, who proved in [7] that in

fact

$$\lim_{k\to\infty}\frac{1}{k}\log R_k^-(\omega)=h \tag{1.1}$$

almost surely.

Wyner and Ziv's definition of R_k^- uses the past and future in an asymmetric fashion, but in [7] Ornstein and Weiss also define a more symmetric quantity, and prove a related limit result, which they later generalized in [8] to the case of $\mathbb{Z}^d$ processes, or so-called random fields. These are families of random variables $\{X_u\}_{u\in\mathbb{Z}^d}$ which are stationary with respect to "time shifts" in $\mathbb{Z}^d$. In this setting, the patterns we are trying to match, and the sets in which we look for recurrences of patterns, are based on the cubes $F_n=[-n;n]^d$ (here and throughout, we denote the integer segment $\{i,i+1\ldots,j\}$ for $i<j$ by $[i;j]$). For a realization $x=(x_u)_{u\in\mathbb{Z}^d}$ of the process, we obtain a "pattern" on the cube F_k, which we denote $x(F_k)$, by coloring each $v\in F_k$ with the color x_v. The first recurrence time of the F_k-pattern $x(F_k)$ is defined to be the first index n for which we observe the pattern $x(F_k)$ centered at some point in F_n other than the origin. To be precise,

$$R_k(\omega)=\min\left\{n\;\middle|\;\begin{array}{c}\text{for some } v\neq 0 \text{ in } [-n;n]^d\\ x_{v+u}=x_u \text{ for every } u\in F_k\end{array}\right\}$$

or $R_k(\omega)=\infty$ if the pattern does not recur. Then according to [8], with probability 1

$$\lim_{k\to\infty}\frac{d}{(2k+1)^d}\log R_k(\omega)=h \tag{1.2}$$

(in [8] there is also a proof for patterns based on the cubes $[0;n]^d$).

Our goal in this paper is to describe several generalizations of these phenomena for processes indexed by other amenable groups, though many of the results are new even for $\mathbb{Z}$-processes

1.2. **Notation and preliminaries.** Recall that a discrete countable group G is amenable if there exists a sequence of finite subsets $\{F_n\}$ of G which are asymptotically invariant, ie for all $g\in G$,

$$\lim_{n\to\infty}\frac{|F_n\,\Delta\,gF_n|}{|F_n|}=0.$$

For example, the sequence $\left\{[-n;n]^d\right\}_{n=1}^{\infty}$ in $\mathbb{Z}^d$ has this property. Such sequences are called Følner sequences. Asymptotic invariance is equivalent to the following statement: for every finite $K\subseteq G$, it holds that

$$\lim_{n\to\infty}\frac{1}{|F_n|}\left|\{f\in F_n\;:\;Kf\subseteq F_n\}\right|=1.$$

There are other equivalent definitions of amenability, some of which have a more analytic flavor, but these will not interest us here; see eg [2].

For G a discrete amenable group we consider an ergodic process $\{X_g\}_{g\in G}$ taking values in a finite set Σ, defined on a standard probability space $(\Omega, \mathcal{F}, P)$. We will always assume that the process arises from an ergodic measure preserving free left action of G on Ω, denoted $(g,\omega)\mapsto g\omega$, and from a measurable function $X:\Omega\to\Sigma$ such that $X_g(\omega)=X(g\omega)$. We denote by h the entropy of the process.

It is convenient to associate the elements $\omega\in\Omega$ with the Σ-colorings of G they induce: $g\mapsto X_g(\omega)$. We will often denote elements of Ω by the letter x to emphasize this point of view, and write $x(g)$ for $X_g(x)$. For $x\in\Omega$ and $E\subseteq G$ we denote by $x(E)$ the coloring of E by colors from Σ given by $g\mapsto x(g)$, and let $[x(E)]\in\mathcal{F}$ be the atom

$$[x(E)] = \{x'\in\Omega\,:\,\forall f\in E\;\; x'(f)=x(f)\}.$$

More generally, if $\varphi\in\Sigma^E$ is a coloring of E by Σ, we write $[\varphi]$ for the atom defined by φ

$$[\varphi] = \{x\in\Omega\,:\,\forall f\in E\;\; x(f)=\varphi(f)\}$$

and for $\Phi\subseteq\Sigma^{F_n}$ we set

$$[\Phi]=\bigcup_{\varphi\in\Phi}[\varphi].$$

It is known that the ergodic and entropy theorems hold along certain Følner sequences. The weakest condition known to ensure this is

Definition 1.1. A Følner sequence $\{F_n\}$ in G is *tempered* if for some constant C independent of n, $|(\bigcup_{k<n}F_k^{-1})F_n|\le C|F_n|$ for every n.

(We extend the group operations to sets in the usual manner, so $AB=\{ab\,:\,a\in A\,,\,b\in B\}$, etc).

Throughout this paper, $\{F_n\}$ will denote a tempered Følner sequence which also satisfies $|F_n|\ge n$. It is not hard to see that if G is an infinite group any Følner sequence in G has a tempered subsequence satisfying this.

For a tempered Følner sequence, the ergodic theorem states that for every $f\in L^1(\Omega,\mathcal{F},P)$,

$$\lim_{n\to\infty}\frac{1}{|F_n|}\sum_{g\in F_n}f(g\omega)=\int f\,dP$$

almost everywhere and in L^1.

The entropy of a process $\{X_h\}_{g\in G}$ is defined as in the case of $\mathbb{Z}$-processes by

$$h=\lim_{k\to\infty}\frac{1}{|F_k|}H(\{X_g\}_{g\in F_k})$$

where $H(\cdot)$ is the usual entropy function defined for any finite-valued random variable. For a discussion of entropy theory of amenable processes see [11].

For such processes, the Shannon-McMillan-Breiman (SMB) theorem states that for $\{F_n\}$ a tempered sequence,

$$\frac{1}{|F_n|}\log P([x(F_n)]) \to h$$

for almost every $x \in \Omega$. In this generality, these theorems are due to E. Lindenstrauss [5]; see also [11] and [6].

1.3. **Return times.** In the return-time theorems stated above for $\mathbb{Z}$ and $\mathbb{Z}^d$ processes a single Følner sequence $\{F_n\}$ (segments or cubes) was used both to define the "shape" of the patterns whose repetitions we are looking for and as the shape of the domain in which we seek these repetitions. There is no reason for the same sequence to play both these roles, and it will be useful to separate them. In our formulation of the problem we will consider two sequences of finite subsets of G: $\{F_n\}$ (F for Følner) which will serve as the shape of the patterns, and $\{W_n\}$ (W for window) which will serve as the set within which we search for repetitions. Throughout, $\{F_n\}$ will be a tempered Følner sequence satisfying $|F_n| \geq n$. The properties we demand of $\{W_n\}$ will vary.

We will be interested in the recurrence of F_k-patterns in the windows W_n. We say that $x(F_k)$ *repeats* at a point $f \in G$ if $fx \in [x(F_k)]$, ie if $fx(g) = x(g)$ for every $g \in F_k$. If $x(F_k)$ repeats at f, we say that the repetition is disjoint from F_k (or from $x(F_k)$) if $F_k \cap F_k f = \emptyset$.

Define

$$R_k(x) = R_k^{(F,W)}(x) = \min\left\{ n \;\middle|\; \begin{array}{c}\text{there exists a disjoint}\\ \text{repetition of } x(F_k) \text{ in } W_n\end{array} \right\}$$

or $R_k(x) = \infty$ if the set on the right in empty.

Remarks.

(1) $R_k(x)$ is the first "return index" of x to $[x(F_k)]$, *excluding* recurrences which intersect the original pattern $x(F_k)$. In this respect our definition differs from the definition for the case $G = \mathbb{Z}^d$ from [8] described in the previous section, where it is not required that the repetition be disjoint from the original occurrence of the pattern.

(2) We require only that the "center" of the repetition f to be in W_n. We do not require the whole of the repetition to be in W_n, ie we do not require $F_k f \subseteq W_n$. Note that in the known cases of cubes, $F_n, W_n = [-n; n]^d \subseteq \mathbb{Z}^d$, requiring that the repetition of $x(F_k)$ be completely contained in W_{R_k} does not change the asymptotic behavior of R_k. In general, however, the definitions are not equivalent. Indeed it is not necessary that any translate of F_k be contained in any W_n.

(3) In what follows we will be working with several different sequences and then the notation $R_k^{(F,W)}$ will be used to emphasize the dependence on $\{F_n\}$ and $\{W_n\}$.

(4) If $\{F_n\}$ is any sequence (not necessarily Følner) and $\{W_n\}$ is a tempered Følner sequence then the ergodic theorem implies that $R_k^{(F,W)}(x)$ is finite almost surely. In other cases it is not clear that $R_k < \infty$, and for general window sequence this may not be the case.

R_k is not a very "stable" quantity, and even when finite we cannot hope for its growth rate to provide any significant information about the process, since it strongly reflects the indexing of the sequence $\{W_n\}$. For any G-process and choice of $\{F_n\}$ and $\{W_n\}$, the value of R_k may be arbitrarily increased simply by allowing repetitions in the W_n's, or decreased by thinning the window sequence out. A better measure of return time is the volume (=size) of the window set in which the repetition is observed. We therefore define

$$T_k = T_k^{(F,W)} = |W_{R_k^{(F,W)}}|$$

or $T_k = \infty$ if $R_k = \infty$, and set

$$T^* = T^*_{(F,W)} = \limsup_{k\to\infty} \frac{1}{|F_k|} \log T_k^{(F,W)}$$

$$T_* = T_*^{(F,W)} = \liminf_{k\to\infty} \frac{1}{|F_k|} \log T_k^{(F,W)}.$$

The aforementioned result for $\mathbb{Z}^d$ can be given in terms of T_k, since for $W_n = [-n;n]^d$, the size of W_{R_k} is $(2R_k+1)^d$; the result then is equivalent to $T^* = T_* = h$.

One half of the $\mathbb{Z}^d$ return-time theorem is true in the following very general form, which provides new information even about $\mathbb{Z}$ processes:

Theorem 1.2. *Let G be a countable amenable group and $\{X_g\}_{g\in G}$ a finite-valued ergodic G-process. If $\{F_n\}$ is a tempered Følner sequence in G and $\{W_n\}$ an increasing sequence of finite subsets of G, then for almost every x,*

$$T_*^{(F,W)}(x) = \liminf_{k\to\infty} \frac{1}{|F_k|} \log T_k^{(F,W)}(x) \geq h.$$

One might hope that the bound $T^* \leq h$ is also true; if it were, we would have $\lim \frac{1}{|F_k|} \log T_k = h$. This need not be true for general window sequences. The asymmetry between T_* and T^*, which does not appear in the special cases examined in [8], is not due to some shortcoming of our methods, but is inherent in the problem. To begin with, in order for the upper bound to hold, at the very least some growth condition must be imposed on $\{W_n\}$. For example if we set $F_n = W_n = [0; 2^{|F_{n-1}|}] \subseteq \mathbb{Z}$ then $|F_{n+1}| > 2^{|F_n|}$ so $\frac{1}{|F_n|} \log T_k \geq 1$ is always true, even if $h < 1$. We therefore cannot hope for a good upper bound on the growth rate of T_k if we do not restrict the growth of the window sequence $\{W_n\}$.

Even if the sequence $\{W_n\}$ grows slowly, the inequality $T^* \leq h$ need not hold. For a simple example, consider the process of alternating sequences of 0s and 1s with equal probabilities for observing 0 or 1 at any given place, so the possible realizations are $\ldots 0101010 \ldots$ and $\ldots 1010101 \ldots$. This process is clearly ergodic and has entropy 0. If we take $F_n = [-n; n]$ and W_n to be sets containing only odd integers, eg $W_n = [-n, n] \cap (2\mathbb{Z}+1)$, then $R_k = \infty$ for every realization. In this example $\{W_n\}$ is of course not a Følner sequence; in section 2.7 we show that the upper bound may fail even when $\{W_n\}$ are "almost" segments and possess extremely good properties which are known to be sufficient for the ergodic and SMB theorems to hold.

In order to state our results about T^* we will require $\{W_n\}$ to have some additional combinatorial properties, which will play a role, in various forms, also in the density theorems stated below.

Definition 1.3. Let $\{W_n\}$ be some sequence of finite subsets of G. A sequence of right-translates $\{W_{n(i)} f_i\}_{i=1}^{I}$ ($f_i \in G$) is called *incremental* if

(1) $n(1) \geq n(2) \geq \ldots \geq n(I)$.
(2) $f_i \notin \cup_{j<i} W_{n(j)} f_j$ for $1 \leq i \leq I$.

Definition 1.4. Let $\{W_n\}$ be a sequence of finite subsets of G with $1_G \in W_n$. $\{W_n\}$ is *filling with constant* C (or C-filling) if for every incremental sequence $\{W_{n(i)} f_i\}$ it holds that $|\cup W_{n(i)} f_i| \geq C \sum |W_{n(i)} f_i|$. A set W is individually C-filling if for every incremental sequence $\{W f_i\}$ it holds that $|\cup W f_i| \geq C \sum |W f_i|$.

The statement that $\{W_n\}$ is filling means, informally, this: suppose one tries to cover G with translates of the W_n, using a "greedy" algorithm, that is, putting down some translate of a W_n, then another (possibly smaller) W_k over a point not yet covered, etc. Note that we allow the different translates to intersect. Then if you stop after finitely many steps, the size of what has in fact been covered will be proportional to the total of the sizes of the W_n's used up until that stage.

A fixed set W is filling with constant $\frac{1}{|W|}$ because in an incremental sequence each point is covered at most $|W|$ times, but generally sequences $\{W_n\}$ containing infinitely many different sets will not be filling, as demonstrated by the sequence $\{[0; n]\}_{n=1}^{\infty}$ in $\mathbb{Z}$ (though not filling, this is a tempered Følner sequence). On the other hand one can verify that $\{[-n; n]\}_{n=1}^{\infty}$ is filling in $\mathbb{Z}$. Still another example is the sequence $\{[-n; n] \cap 2\mathbb{Z}\}$ in $\mathbb{Z}$ which is filling but not a Følner sequence.

From these examples it should be clear that the notion of being filling is rather delicate, and not related directly to amenability. In particular, note that while a Følner sequence is in some sense a large subset of the group, a filling sequence can be small, eg contained in a proper subgroup.

Definition 1.5. An increasing sequence $\{W_n\}$ of finite subsets of G with $1_G \in W_n$ is *quasi-filling* if each W_n is C_n-filling and $\sum \frac{|W_n|^{-\alpha}}{C_n} < \infty$ for every $\alpha > 0$.

Note that in general filling sequences need not be quasi filling.

Theorem 1.6. *Let G be a countable amenable group and $\{X_g\}_{g\in G}$ a finite-valued ergodic G-process. If $\{F_n\}$ is a tempered Følner sequence in G and $\{W_n\}$ either a filling or a quasi-filling sequence such that $\lim_n \frac{\log|W_n|}{\log|W_{n-1}|} = 1$, then*

$$T^* = \limsup_{k\to\infty} \frac{1}{|F_k|} \log T_k^{(F,W)} \le h$$

almost surely.

Corollary 1.7. *If $\{F_n\}, \{W_n\}$ satisfy the conditions of theorem 1.6 and 4.3, then*

$$\lim_{k\to\infty} \frac{1}{|F_k|} \log T_k^{(F,W)} = h$$

almost surely.

It is interesting to note that in [8], the return times theorem was proved also for the case $F_n = W_n = [0;n]^d \subseteq \mathbb{Z}^d$, which is not covered by our results, because $\{[0;n]^d\}$ is not (quasi) filling. We will not discuss this here.

1.4. **Recurrence densities.** As we have seen, if $\{W_n\}$ grows too swiftly it may not be possible for $T^* \le h$ to hold, because the first W_n within which we observe $x(F_k)$ may be too large compared to F_k. To compensate for this, one expects that the number of times the pattern $x(F_k)$ appears in W_n will grow along with $|W_n|$.

Another way of looking at it is this: Suppose $\{F_n\}, \{W_n\}$ are sequences such that the return times theorem of the last section holds: $\frac{1}{|F_k|}\log T_k^{(F,W)} \to h$. If $\{W_n\}$ grows slowly enough, then one sees from the return time theorem that the density of recurrences of $x(F_k)$ in $W_{R_k(x)-1}$ is approximately $2^{-h|F_k|}$, since only the central copy of the pattern exists. Now suppose in addition that $\{W_n\}$ is itself a tempered Følner sequence; since by the SMB theorem the probability of the atom $[x(F_k)]$ is approximately $2^{-h|F_k|}$, the ergodic theorem tells us that for large enough n (depending on k and x) the frequency with which $x(F_k)$ repeats in $x(W_n)$ is approximately equal to the probability of the atom $[x(F_k)]$. So we see that for large k, the density of recurrences of $x(F_k)$ in $x(W_n)$ is close to $2^{-h|F_k|}$ both when n is small (but large enough for the question to be meaningful) and for large enough n. However, neither the return-time theorem nor the ergodic theorem give us any information about what happens for intermediate values of n.

We define the recurrence frequency of $x(F_k)$ in $x(F_n)$ to be

$$U_{k,n}(x) = U_{k,n}^{(F,W)}(x) = \frac{1}{|W_n|} \max\left\{ |E| \;\middle|\; \begin{array}{c} 1_G \in E \subseteq W_n\text{, the collection} \\ \{F_k f\}_{f\in E} \text{ is pairwise disjoint,} \\ \text{and } x(F_k) \text{ repeats at } f \end{array} \right\}$$

and

$$U^* = U^{(F,W)*} = \limsup_{k\to\infty} \sup_n -\frac{1}{|F_k|} \log U_{k,n}^{(F,W)}$$
$$U_* = U_*^{(F,W)} = \liminf_{k\to\infty} \inf_n -\frac{1}{|F_k|} \log(U_{k,n}^{(F,W)} - \frac{1}{|W_n|}).$$

Remarks:

(1) The quantity $U_{k,n}$ was defined as the (maximal) density of *disjoint* repetitions of $x(F_k)$ in W_n, one of which is the original pattern. One may attempt to define $U_{k,n}$ differently, eg by counting every repetition or counting every repetition disjoint from the original pattern (but not mutually disjoint). These variations are discussed in more detail in sections 3.1 and 4.5.

(2) The reason for the correction factor $-\frac{1}{|W_n|}$ in the definition of U_* is that our goal is to prove $U_* \geq h$, but in our definition of $U_{k,n}$ we allowed the "central" copy of $x(F_k)$ to be counted. This being the case, we cannot expect the observed frequency $U_{k,n}$ to drop below $2^{-h|F_k|}$ if the size $|W_n|$ is not on the order of $2^{h|F_k|}$.

As in the case of T^*, in order to prove $U_* = U^* = h$ we need some assumptions about the combinatorial properties of the window sequence $\{W_n\}$. For the bound $U^* \leq h$, it is enough that $\{W_n\}$ be filling (definition 1.4 above):

Theorem 1.8. *Let G be a countable amenable group and $\{X_g\}_{g\in G}$ a finite-valued ergodic G-process. Let $\{F_n\}$ be a tempered Følner sequence in G. If $\{W_n\}$ is a filling sequence, then*

$$U^* = \limsup_{k\to\infty} \sup_n -\frac{1}{|F_k|} \log U_{k,n}^{(F,W)} \leq h$$

almost surely.

See also theorem 3.3 below for a slightly weaker version which applies to quasi-filling sequences.

For the lower bound $U_* \geq h$ we need an even stronger notion of "filling up space":

Definition 1.9. An increasing sequence $\{W_n\}$ of finite subsets of G with $1_G \in W_n$ is said to be *incompressible with constant C* (or C-incompressible) if for any incremental sequence $\{W_{n(i)}f_i\}$, the number of the sets $W_{n(i)}f_i$ containing 1_G is at most C. A finite set $W \subseteq G$ containing 1_G is *individually C-incompressible* if for any incremental sequence $\{Wf_i\}$, the number of the sets Wf_i containing 1_G is at most C.

Clearly the condition that 1_G be in at most C sets is equivalent to requiring that every $g \in G$ is contained in at most C sets. It follows that a C-incompressible sequence is $\frac{1}{C}$-filling.

Any finite set W containing 1_G is $|W|$-incompressible. Ascending sequences of subgroups, and ascending centered cubes in $\mathbb{Z}^d$, are examples of incompressible sequences (see also section 2).

We also need the following:

Definition 1.10. Let $\{F_n\}, \{W_n\}$ be sequences of finite subsets of G. A sequence $\{Y_n\}$ of finite subsets of G is called an *interpolation sequence for* $\{F_n\}, \{W_n\}$ if

(1) $\{Y_n\}$ is filling.
(2) There exists a constant C such that if $|W_n| \geq |Y_k|$ then $|Y_k W_n| \leq C|W_n|$.
(3) For every pair of real numbers $0 \leq \alpha < \beta$, for every large enough n there is an index k such that $2^{\alpha|F_n|} \leq |Y_k| \leq 2^{\beta|F_n|}$.

In the case where $F_n = W_n$, we say simply that $\{Y_k\}$ is an interpolation sequence for $\{F_n\}$.

Note that $\{F_n\}$ appears only in condition (3) of the definition, and will be satisfied automatically (for any sequence $\{F_n\}$) if $\lim_n \frac{\log|Y_n|}{\log|Y_{n-1}|} = 1$.

As an example, one may consider the sequence $\{[-n;n]\}$ of intervals in $\mathbb{Z}$, which is its own interpolation sequence. In fact if G contains an element of infinite order, and we identify the infinite-cyclic group it generates with $\mathbb{Z}$, then segments in this subgroup can often serve as an interpolation sequence for a suitably chosen window sequence in G. We come back to this in section 2.6.

Theorem 1.11. *Let G be a countable amenable group and $\{X_g\}_{g\in G}$ a finite-valued ergodic G-process. Let $\{F_n\}$ be a tempered Følner sequence in G. Let $\{W_n\}$ be an incompressible sequence. If there exists an interpolation sequence for $\{F_n\}, \{W_n\}$, then*

$$U_*^{(F,W)} = \liminf_{k\to\infty} \inf_n -\frac{1}{|F_k|} \log(U_{k,n}^{(F,W)} - \frac{1}{|W_n|}) \geq h$$

almost surely.

Since $T_* \geq h$ almost surely, we have $U_* = \liminf_k \inf_n -\frac{1}{|F_k|} \log U_{k,R_k}$ (ie the correction term may be dropped), so

Corollary 1.12. *If $\{F_n\}$,$\{W_n\}$ are as in theorem 1.11, then*

$$\lim_k -\frac{1}{|F_k|} \log U_{k,R_k}^{(F,W)} = h$$

almost surely.

A slightly weaker condition than incompressibility is the following:

Definition 1.13. Let $\{W_n\}$ be an increasing sequence of finite subsets of G containing 1_G, and suppose that each W_n is individually C-incompressible (for the same C). $\{W_n\}$ is *quasi-incompressible* if for every $\lambda > 0$ and large enough k, for every $n > k$ the W_k boundary of W_n is at most a $|W_k|^{-\lambda}$

fraction of W_n, ie

$$\frac{1}{|W_n|}\{f \in W_n \,:\, W_k f \not\subseteq W_n\} \leq |W_k|^{-\lambda}.$$

Note that with this definition, an incompressible sequence may not be quasi-incompressible.

Theorem 1.14. *Let G be a countable amenable group and $\{X_g\}_{g\in G}$ a finite-valued ergodic G-process. Let $\{F_n\}$ be a tempered Følner sequence and $\{W_n\}$ a quasi-incompressible sequence in G. If there exists an interpolation sequence for $\{F_n\},\{W_n\}$, then for almost every x,*

$$U_*^{(F,F)}(x) = \liminf_{k\to\infty} \inf_n -\frac{1}{|F_k|}\log(U_{k,n}^{(F,F)}(x) - \frac{1}{|F_n|}) \geq h.$$

Remark. We have not required in any of the theorems that $\{W_n\}$ be a Følner sequence. In fact one may verify that if $H < G$ and $\{W_n\}$ is a sequence of subsets of H satisfying one of the various filling properties described above (except for the property of the existence of an interpolation sequence), then the property continues to hold when $\{W_n\}$ is viewed as a sequence of subsets of G. This means that the recurrence phenomena we have described can often be observed when one counts repetitions in a subgroup of G. We will come back to this in more detail in section 2.6 below.

The rest of this paper is organized as follows. In section 2 we discuss various ways to construct good window sequences and classes of groups in which there exist good window sequences, with respect to which the recurrence phenomena hold. We also give an example showing that other plausible combinatorial properties of $\{W_n\}$ do not ensure that the upper bounds for T^* and U^* hold. In section 3 we prove the upper bounds $T^*, U^* \leq h$. In section 4 we prove the lower bound $T_* \geq h$, whose proof has a different flavor than the others, and $U_* \geq h$, which depends strongly on the lower bound $T_* \geq h$. Finally in section 5 we state some open questions.

Acknowledgement. The results in this paper are part of the author's MA thesis, conducted under the guidance of Prof. B. Weiss, whom I would like to thank for all his patience and good advice.

2. Constructions, examples and counterexamples

In this section we will examine the problem of the existence of good window sets. We consider two questions: When do good window sets exist at all, and when do "large" window sets exist. By large we mean window sequences which are themselves Følner sequences. We also give an example which shows that, while the filling properties may not be strictly necessary for our theorems to hold, other combinatorial conditions on the window sequence, which are known to be sufficient for other pointwise theorems, are not sufficient for upper bounds on return times and densities.

2.1. **General constructions.** Before going into concrete examples we examine how good window and interpolation sequences can be constructed from other sequences.

Lemma 2.1. *Suppose $\{W'_n\}$ is C'-incompressible in G' and $\{W''_n\}$ is C''-incompressible in G''. Set $W_n = W'_n \times W''_n$. The $\{W_n\}$ is incompressible in $G = G' \times G''$.*

Proof. Clearly $\{W_n\}$ is increasing and $1_G \in W_n$ for all n. Suppose $\{W_{n(i)} g_i{}_{i=1}^I\}$ is incremental and $1_G \in \cap_i W_{n(i)} g_i$; we must show that I is bounded. Write $g_i = (g'_i, g''_i)$. Since $1_{G'} \in \cap_i W'_{n(i)} g'_i$ and $1_{G''} \in \cap_i W''_{n(i)} g''_i$, it suffices to show that there is a C such that if $I > C$ then either there is a subsequence $\{i(1), \ldots, i(C'+1)\} \subseteq \{1, \ldots, I\}$ with $\{W'_{n(i(j))} g'_{i(j)}\}_{i=1}^{C'+1}$ incremental or there is a subsequence $\{i(1), \ldots, i(C''+1)\} \subseteq \{1, \ldots, I\}$ with $\{W'_{n(i(j))} g'_{i(j)}\}_{i=1}^{C''+1}$ incremental.

We prove this claim by induction on $C' + C''$. To be precise, we claim that for each pair C', C'' there is a number $\alpha(C', C'')$ such that if $\{W_{n(i)} g_i\}_{i=1}^I$ is an incremental sequence in G of length $I > \alpha(C', C'')$ then either its projection on the first coordinate has an incremental subsequence of length greater than C' or its projection on the second coordinate has an incremental subsequence of length C''. Clearly

$$\alpha(C', 1) = C' \; \alpha(1, C'') = C'' \; \alpha(1, 1) = 1.$$

For the induction step, let $\left\{W_{n(i)} g_i\right\}_{i=1}^I$ be incremental in G. Write

$$U_1 = \left\{1 < i \le I \, : \, g'_i \notin W'_{n(1)} g'_1\right\} \; , \; U_2 = \left\{1 < i \le I \, : \, g''_i \notin W''_{n(1)} g''_1\right\}$$

clearly $U_1 \cup U_2 = \{2, \ldots, I\}$. If $|U_1| > \alpha(C'-1, C'')$ or $|U_2| > \alpha(C', C''-1)$ then $\left\{W_{n(i)} g_i\right\}$ has a subsequence of the type we desire. But one of these relations will certainly hold if $|U_1 \cup U_2| > \alpha(C'-1, C'') + \alpha(C', C''-1)$, so we can set

$$\alpha(C', C'') = \alpha(C'-1, C'') + \alpha(C', C''-1) + 1. \qquad \square$$

The point here is that we obtain a "large" incompressible sequence in $G' \times G''$. "Small" sequences exist since an incompressible sequence in G' (or G'') is incompressible in $G' \times G''$ with the natural embedding. Note that if $\{W'_n\}, \{W''_n\}$ are Følner sequences in G', G'' respectively then $\{W'_n \times W''_n\}$ is Følner in $G' \times G''$.

In [10] it was shown how certain properties of Følner sequences can be pulled up through exact sequences. We rematk that we do not know whether the existence of filling or incompressible Følner sequences in K and G/K for a normal subgroup K of G implies that such sequences exist in G, although it seems unlikely that this is the case (again, the sequence in K has the same good filling properties in G, but is not Følner).

The requirement that $\{W_n\}$ be C-filling (or C-incompressible) is a property of the entire sequence, and in particular implies that each W_n is C-filling (C-incompressible) as an individual set. The

existence of a constant C such that each W_n is C-filling (C-incompressible) does not imply that the same about the sequence. One can, however, sometimes pass to a subsequence of the $\{W_n\}$ which will be filling.

Lemma 2.2. *Suppose $\{W_n\}$ is an increasing sequence of finite subsets of G with $1_G \in W_n$ such that*

1. *$\{W_n\}$ is a Følner sequence.*
2. *For some constant C, each W_n is individually C-filling.*

Then there exists a subsequence $\left\{W_{n(i)}\right\}$ which is filling.

Proof. Since $\{W_n\}$ is a Følner sequence, we may select a subsequence $\{W_{n(i)}\}$ with the property that

$$\frac{1}{|W_{n(i)}|}\left|\left\{g \in W_{n(i)} \,:\, W_{n(i-1)}g \subseteq W_{n(i)}\right\}\right| \geq \frac{C}{2}.$$

Write $W'_i = W_{n(i)}$; we claim that $\{W'_i\}$ is filling. Suppose that $\{W'_{i(j)}f_j\}_{j=1}^{J}$ is an incremental sequence. Let $I(1) > \ldots > I(N)$ be all the values of $i(1), \ldots, i(J)$, and let $J_k = \{1 \leq j \leq J \,:\, I(j) = k\}$. Since $W_{I(1)}$ is C-filling, we have

$$|\bigcup_{j\in J_1} W_{I(1)}f_j| \geq C\sum_{j\in J_1}|W_{I_1}f_j|.$$

Let

$$E_1 = \{g \in \bigcup_{j\in J_1} W_{I(1)}f_j \,:\, W_{I(2)}g \subseteq \bigcup_{j\in J_1} W_{I(1)}f_j\}$$

then

$$\begin{aligned}
|E_1| &\geq |\bigcup_{j\in J_1} W_{I(1)}f_j| - |\{g \in \bigcup_{j\in J_1} W_{I(1)}f_j \,:\, W_{I(2)}g \not\subseteq \bigcup_{j\in J_1} W_{I(1)}f_j\}| \\
&\geq C\sum_{j\in J_1}|W_{I(1)}f_j| - \sum_{j\in J_1}\frac{C}{2}|W_{I(1)}f_j| \\
&= \frac{C}{2}\sum_{j\in J_1}|W_{I(1)}f_j|
\end{aligned}$$

so

$$|\bigcup_{j\in J_1} W_{I(1)}f_j \setminus \bigcup_{j\in J_2\cup\ldots\cup J_N} W_{i(j)}f_j| \geq |E_1| \geq \frac{C}{2}\sum_{j\in J_1}|W_{I_1}f_j|$$

which means that the fraction of $\bigcup_{j\in J_1} W_{I(1)}f_j$ which is not contained in any of the "lower levels" is at least $\frac{C}{2}\sum_{j\in J_1}|W_{I(1)}f_j|$. Repeating this argument for $I(1),\dots,I(N)$ we obtain

$$|\bigcup_{1\le j\le J} W_{i(j)}f_j| \ge \frac{C}{2}\sum_{1\le j\le J}|W_{i(j)}f_j|. \qquad \square$$

The following is proved similarly:

Lemma 2.3. *Suppose $\{W_n\}$ is an increasing sequence of finite subsets of G with $1_G \in W_n$ such that*

(1) *$\{W_n\}$ is a Følner sequence.*
(2) *For some constant C, each W_n is C-incompressible.*

Then there exists a subsequence $\{W_{n(i)}\}$ which is quasi-incompressible .

The last construction in this section deals with the existence of interpolation sequences.

Lemma 2.4. *Let $\{F_n\}$ be an increasing tempered Følner sequence. Suppose $\{Y_n\}$ is increasing and filling. Then there exists a subsequence $\{F_{n(i)}\}$ of $\{F_n\}$ such that the sequence Y_n is an interpolation sequence for $\{F_{n(i)}\}$.*

For the sequence $\{Y_n\}$ to be useful, $\{F_n\}$ must fulfill the hypothesis of theorem 1.11 and so $\{F_n\}$ must also be incompressible or quasi-incompressible; but this is not necessary for the proof of the lemma.

Proof. Let $\{[\alpha_i,\beta_i]\}_{i=1}^{\infty}$ be an enumeration of the rational intervals, $0\le\alpha_i<\beta_i$. We select $F_{n(k)}$ and sets $Z_{i,k}=Y_{m(i,k)}$ $(1\le i<k)$ inductively in k, as follows: Given k and assuming that we have defined $F_{n(j)}$ and $Z_{i,j}$ for every $1\le i<j<k$, we select $F_{n(k)}$ and $Z_{i,k}$ for $1\le i<k$ so that

(a) $Z_{i,k}$ satisfies $2^{\alpha_i|F_{n(k-1)}|}\le|Z_{i,k}|\le 2^{\beta_i|F_{n(k-1)}|}$.
(b) $F_{n(k)}$ is large enough that for all $i<k$ it holds that $|F_{n(k)}|<2^{\alpha_i|F_{n(k)}|}$.
(c) $|(\bigcup_{i<k-1} Z_{i,k-1})F_{n(k)}|\le 2|F_{n(k)}|$.
(d) $n(k)$ is large enough so that for $1\le i\le k$ and every $n>n(k)$ there exists m such that $2^{\alpha_i|F_n|}<|Y_m|<2^{\beta_i|F_n|}$.

One then verifies that $\{Z_{i,k}\}$ is an interpolation sequence for $\{F_{n(k)}\}$. Briefly, (a) ensures that (3) of the definition of interpolation sequences holds, while (b) and (c) ensure condition (2), and (d) makes it possible to continue the construction ((a) is possible for k because for $j<k$ we chose $W_{n(j)}$ to satisfy (d)). $\square$

2.2. **The groups $\mathbb{Z}$ and $\mathbb{Z}^d$.** In the case of $\mathbb{Z}$, any increasing sequence of symmetric intervals $\{I_n\}$ is 2-incompressible. This is simple to verify.

Next, $\mathbb{Z}^d$ is the product of $\mathbb{Z}$ with itself d times, we know from lemma 2.1 that incompressible sequences exist. In fact,

Proposition 2.5. *If $\{I_n^{(i)}\}_{n=1}^{\infty}$ are increasing sequences of symmetric intervals, $i = 1, \ldots, d$, then the sets $W_n = I_n^{(1)} \times \ldots \times I_n^{(d)}$ form an 2^d-incompressible sequence.*

Lemma 2.1 gives a poorer estimate for the constant. That the constant 2^d is correct (and optimal) for such sequences can be established directly.

We will call sets of the form in the proposition boxes. If $\{W_n\}$ is an increasing symmetric sequence of boxes, then it is easy to see that for any n we can find boxes $W_n = Y_{n,0} \subseteq \ldots \subseteq Y_{n,k(n)} = W_{n+1}$ such that $|Y_{n,i+1}|/|Y_{n,i}| \leq 2$. Ordering all the $Y_{n,i}$ in a single increasing sequence $\{Y_m\}$, we obtain a filling sequence which grows at most exponentially. It is easy to verify that $|W_n + Y_m| \leq 2^d|W_n|$ whenever $|W_n| \geq |Y_m|$; so $\{Y_m\}$ is an interpolation sequence for $\{W_n\}$. This gives

Proposition 2.6. *For $G = \mathbb{Z}^d$, any increasing sequence $\{W_n\}$ of boxes and any tempered Følner sequence $\{F_n\}$ (and in particular $F_n = W_n$), $T_*^{(F,W)} \geq h$ and $U_*^{(F,W)} = U^*_{(F,W)} = h$. If $\lim_n \frac{\log|W_{n+1}|}{\log|W_n|} = 1$ then $T_*^{(F,W)} = T^*_{(F,W)} = h$.*

This is essentially a reformulation of the results from [8].

One can show that certain increasing sequences of symmetric convex sets in $\mathbb{Z}^d$ are incompressible as well.

2.3. **The group $\mathbb{Z}^\infty$.** Consider the group $\mathbb{Z}^\infty = \oplus_{n=1}^{\infty}\mathbb{Z}$. Identify the subgroups $\mathbb{Z}^d \times \{(0, 0 \ldots)\} \subseteq \mathbb{Z}^\infty$ with $\mathbb{Z}^d$ in the obvious way.

Proposition 2.7. *Let $A \subseteq \mathbb{Z}^d$ have full dimension (ie $\langle A\rangle$ is of abelian rank d). If A is C-incompressible then $C > d$.*

Proof. For a set $A \subseteq \mathbb{Q}^d$, we say that $u \in \mathbb{Q}^d$ is an extreme point of A if it is an extreme point of $\mathrm{conv}_{\mathbb{Q}} A$.

We prove the following: If $A \subseteq \mathbb{Q}^d$ is finite and has $\dim\langle A\rangle = d$, and 0 is an extreme point of A, then there exists a set $\{u_1, \ldots, u_{d+1}\} \subseteq -A$ such that $\{A + u_i\}_{i=1}^{d+1}$ is incremental. In this case, clearly $0 \in \cap(A + u_i)$, so A cannot be C-incompressible for $C \leq d$.

The proof is by induction on d. For $d = 1$ the claim is trivial: take $a_1 = 0$ and $a_2 \in A \setminus \{0\}$.

Now suppose we have proved it for $d - 1$. Let $A \subseteq \mathbb{Q}^d$ be as in the claim. Set $E = \mathrm{ext}\,\mathrm{conv}\, A$, so by assumption $0 \in E$. Since A is d-dimensional, so is E. Select a linearly independent set $a_1, \ldots, a_{d-1} \in E \setminus \{0\}$. Write $V = \mathrm{span}_{\mathbb{Q}}\{a_1, \ldots, a_{d-1}\}$, and assume $\{a_i\}$ was selected so that A is on one side of the subspace V, ie there is a linear functional Λ with $\Lambda(A) \subseteq \mathbb{Q}^+$ and $V = \ker\Lambda$.

Consider $A' = A \cap V$. This is a finite set of full dimension in V containing 0 as an extreme point. Therefore there exist $u_1, \ldots, u_d \in -A'$ with $\{A' + u_i\}_{i=1}^d$ incremental. Clearly $\{A + u_i\}_{i=1}^d$ is also incremental. Now we need only note that for any $u \in A \setminus A'$, we have that $-u \notin \cup_{i=1}^d A + u_i$ because $\Lambda(u) < 0$ whereas $\Lambda(a + u_i) = \Lambda(a) + 0 \geq 0$ for all $a \in A$ and $i = 1, \ldots, d$. Thus $\{u_1, u_2, \ldots, u_d, -u\}$ is the set we are looking for. This completes the induction step.

The proposition now follows from the fact that it holds for A iff it holds for every translate of A, and we can always translate A so that 0 is an extreme point. □

Corollary 2.8. *There are no incompressible Følner sequences in $\mathbb{Z}^\infty$.*

Proof. If $\{F_n\}$ is a Følner sequence in $\mathbb{Z}^\infty$ then for any d, F_n must be eventually d-dimensional. □

It seems likely that there is a similar bound on how filling a d-dimensional set in $\mathbb{Z}^d$ can be, and this would imply that there are no filling Følner sequences in $\mathbb{Z}^\infty$, but we do not have a proof.

On the bright side, there do exist quasi-filling Følner sequences in $\mathbb{Z}^\infty$, as we now demonstrate.

Suppose $I_n^{(i)} \subseteq \mathbb{Z}$ are symmetric segments and $W_n = I_n^{(1)} \times \ldots \times I_n^{(d(n))} \subseteq \mathbb{Z}^\infty$ is an increasing sequence of finite-dimensional boxes. We know that each set W_n is $2^{-d(n)}$-filling in $\mathbb{Z}^{d(n)}$ and it follows that the same is true as subsets of $\mathbb{Z}^\infty$. In order for $\{W_n\}$ to be quasi filling, we must have

$$\sum_n 2^{d(n)} \cdot |W_n|^{-\alpha} < \infty$$

for every $\alpha > 0$. For this it is enough that $\{W_n\}$ grow exponentially and $d(n) = O(\log |W_n|)$. We would also like $\{W_n\}$ to be a tempered Følner sequence. Such a sequence can easily be constructed. For example, we may take

$$W_n = [-2^{n^2}; 2^{n^2}]^{[\log n]}.$$

$\{W_n\}$ is then an increasing quasi-filling Følner sequence. For temperedness it suffices to check that $|W_{n-1}^{-1} W_n| \leq C|W_n|$. We have

$$W_{n-1}^{-1} W_n = [-2^{n^2} - 2^{(n-1)^2}, 2^{n^2} + 2^{(n-1)^2}]^{[\log n]}$$

so

$$\frac{|W_{n-1}^{-1} W_n|}{|W_n|} \leq \left(\frac{2 \cdot (2^{n^2} + \cdot 2^{(n-1)^2}) + 1}{2 \cdot 2^{n^2} + 1} \right)^{[\log n]} \leq$$

$$\leq \left(1 + 2^{-2n+2}\right)^{[\log n]} \to 1.$$

One also readily verifies that $\{W_n\}$ satisfies $\frac{\log |W_{n+1}|}{\log |W_n|} \to 1$. We therefore have

Proposition 2.9. *There exists a tempered Følner sequence $\{W_n\}$ in $\mathbb{Z}^\infty$ such that for any tempered sequence $\{F_n\}$ (and in particular $F_n = W_n$) we have $T_*^{(F,W)} = T^*_{(F,W)} = h$.*

2.4. **Locally finite groups.** G is *locally finite* if every finitely generated subgroup is finite. Since we only consider countable groups this is equivalent to saying that there are finite subgroups $G_1 < G_2 < \ldots$ whose union is all of G.

It is easy to see that such a sequence $\{G_n\}_{n=1}^{\infty}$ is 1-incompressible (and hence 1-filling). Translates of G_i's are just cosets, so if $f \notin G_i g$ then $G_i f \cap G_i g = \emptyset$. Thus if $\{G_{n(i)} g_i\}_{i=1}^{I}$ is incremental, the $G_{n(i)} g_i$ are pairwise disjoint, so every $f \in G$ belongs to at most one translate. Such a sequence $\{G_n\}$ is also clearly a tempered Følner sequence, since every $g \in G$ is eventually a member of G_n for large enough n.

If in addition $\{G_i\}$ grows slowly enough, then $\{G_i\}$ is its own interpolation sequence. Thus

Proposition 2.10. *If $G_1 < G_2 < \ldots$ and $G = \cup G_i$, then for $F_k = W_k = G_k$ we have that $T_* \geq h$ and $U^* \leq h$. If in addition $\frac{\log |G_{i+1}|}{\log |G_i|} \to 1$ then $T_* = T^* = h$ and $U_* = U^* = h$.*

In the case where $\{G_i\}$ grows too quickly to be its own interpolation sequence, we can still show that for a suitably chosen subsequence $\left\{G_{i(k)}\right\}$ there exists an interpolation sequence. This is based on

Theorem. *(P. Hall and C. R. Kulatilaka,* [4]*) Every infinite locally finite group has an infinite abelian subgroup*

To use the theorem we need

Proposition 2.11. *An infinite locally finite abelian group A possesses an increasing sequence of finite subsets $(Y_n)_{n=1}^{\infty}$each of which is indivdually 4-incompressible and $|Y_{n-1}| < |Y_n| \leq 2|Y_{n-1}|$.*

Proof. We can find finite subgroups $1 < \ldots < A_n < \ldots A$ such that A_{n+1}/A_n is cyclic. Thus $A_{n+1} = A'_{n+1} \oplus C_{n+1}$ and $A_n = A'_{n+1} \oplus k_n C_{n+1}$ for C_{n+1} some cyclic group and $k_n \in \mathbb{N}$. Let $Z_{n,i}$ be a 2-incompressible sequence in C_{n+1}; then $Y_{n,i} = A_n \cup (A'_{n+1} \times Z_{n,i})$ is 4-incompressible, and ordering all the $Y_{n,i}$ by inclusion gives the desired sequence. □

Now if G is locally finite, and $A < G$ an infinite abelian group, we can find a slowly-growing sequence (Y_n) of finite subsets of A each of which is 4-incompressible. Thus (Y_n) enjoys the same properties as subsets of G, and we may apply lemma 2.4 to obtain

Proposition 2.12. *For any locally finite group $G = \cup_{n=1}^{\infty} G_n$ there exists a sequence $G_{n(i)}$ such that for $F_i = G_{n(i)}$,*

$$U_*^{(F,F)} = U^*_{(F,F)} = h.$$

We remark that in order to obtain an interpolation sequence in G one does not need the full force of the Hall-Kulatilaka theorem, but can rather use the fact that for every N there is a K such that

every finite group H with $|H| > K$ has an abelian subgroup of size at least N, and therefore infinite locally finite groups have arbitrarily large abelian subgroups. This can be proved by elementary methods.

2.5. **Groups with polynomial growth.** Let G be a finitely generated group and $\Gamma \subseteq G$ a finite set of generators. For convenience we assume throughout that Γ is symmetric, ie that $\Gamma = \Gamma^{-1}$. For every element $g \in G$ define the *length* of g (with respect to Γ) by

$$\ell_\Gamma(g) = \min\{k \,:\, g \in \Gamma^k\}$$

so $\ell_\Gamma(g)$ is the least number of elements of Γ which need be multiplied together to give g.

It is easy to check that

$$d_\Gamma(g_1, g_2) = \ell_\Gamma(g_1^{-1} g_2)$$

defines a left invariant metric on G (this is just the shortest distance metric on G's Cayley graph). The ball of radius n in this metric centered at the unit element of G is

$$B_n^\Gamma = \{g \in G \,:\, \ell_\Gamma(g) \leq n\}.$$

The *growth function* of G (with respect to Γ) is defined by

$$\gamma_\Gamma(n) = |B_n^\Gamma| = \#\{g \in G \,:\, \ell_\Gamma(g) \leq n\}.$$

A finitely generated group G is said to have *polynomial growth* if its growth function γ_Γ is bounded by a polynomial, and *subexponential growth* if $\frac{1}{n}\log \gamma_\Gamma(n) \to 0$. The property of having polynomial/subexponential growth is really a property of the group and does not depend on the particular generating set Γ. To see this, let Γ_1, Γ_2 be two finite generating sets for G as above. Then there is an N for which $\Gamma_1 \subseteq (\Gamma_2)^N$, so $(\Gamma_1)^n \subseteq (\Gamma_2)^{Nn}$. This means that

$$\ell_{\Gamma_1}(g) \leq N \ell_{\Gamma_2}(g)$$

and so

$$\gamma_{\Gamma_1}(n) \leq \gamma_{\Gamma_2}(Nn).$$

From this, and the fact that the roles of Γ_1 and Γ_2 are interchangeable, it follows that γ_{Γ_1} is bounded from above by a polynomial $p(x)$ of degree d iff γ_{Γ_2} is also (though not necessarily the same polynomial), and similarly $\frac{1}{n}\log \gamma_{\Gamma_1}(n) \to 0$ iff $\frac{1}{n}\log \gamma_{\Gamma_2}(n) \to 0$.

Though formally finite groups have polynomial growth, we will assume from here on that our groups are infinite. We also assume that Γ is a fixed finite symmetric generating set for G and write γ for γ_Γ and B_n for B_n^Γ.

One can show using elementary methods that groups with subexponential growth are amenable. Indeed,

Proposition 2.13. *There is a subset $I \subseteq \mathbb{N}$ of density 1 such that $\{B_i\}_{i \in I}$ is a Følner sequence.*

Proof. For suppose that γ grows subexponentially. We will find a subsequence $\{B_{n_k}\}_{k=1}^{\infty}$ of balls which is a Følner sequence.

Write $\alpha(n) = \frac{\gamma(n+1)}{\gamma(n)}$ and for $\varepsilon > 0$ set $I_\varepsilon = \{n \in \mathbb{N} : \alpha(n) \geq 1 + \varepsilon\}$. Since γ grows subexponentially, for every $\varepsilon > 0$, I_ε must have zero density; ie

$$\lim_{n \to \infty} \frac{1}{n+1} |I_\varepsilon \cap [0; n]| = 0.$$

Choose N_k growing very rapidly such that

$$\frac{1}{N_k + 1} |I_{1/k} \cap [0; N_k]| < 1/k.$$

And define $I = \cup_k (I_{1/k} \cap [N_{k-1}; N_k])$. Since if n is such that $\alpha(n) \geq 1 - \frac{1}{k}$ then $gB_n \subseteq B_{n+1}$ and

$$|gB_n \Delta B_n| \leq |B_{n+1} \setminus gB_n| + |B_{n+1} \setminus B_n| \leq \frac{2}{k} |B_n|.$$

We see that $(B_n)_{n \in I}$ is Følner, and clearly I has density 1. □

We will need the following stronger fact:

Theorem. *([1] and [3]) If $G = \langle \Gamma \rangle$ has polynomial growth then there is an integer d and constants c_1, c_2 such that $c_1 n^d \leq \gamma_\Gamma(n) \leq c_2 n^d$.*

This implies that

$$|B_n^{-1} B_n| = |B_n^2| \leq c_2 (2n)^d \leq \frac{2^d c_2}{c_1} |B_n|$$

so the Følner sequence from the proposition is in fact tempered.

We also conclude that every ball B_n is $4^d c_2 / c_1$ incompressible. For suppose that $\{B_n g_i\}_{i=1}^N$ is incremental and $1_G \in \cap_i B_n g_i$. Then $\{B_{n/2} g_i\}$ is a disjoint collection contained in $B_{2n} g_i$ so

$$N \cdot c_1 (\frac{n}{2})^d \leq N \cdot \gamma_{n/2} = \sum_{i=1}^{N} \gamma_{n/2} \leq \gamma_{2n} \leq c_2 (2n)^d$$

and therefore $N \leq \frac{c_2}{c_1} 4^d$.

We thus get

Proposition 2.14. *If G has polynomial growth there is a Følner sequence $\{F_n\}$ such that $\frac{1}{|F_n|} \log T_k^{(F,F)} \to h$ almost surely.*

Using the construction from 2.2 and 2.4 we also get

Proposition 2.15. *If G has polynomial growth there is a Følner sequence $\{F_n\}$ such that $U^*_{(F,F)} = U_*^{(F,F)} = h$ almost surely.*

We also note that by Gromov's theorem [3] every infinite group G with polynomial growth has an element of inifinite order, so the next section applies.

2.6. **Recurrence in subgroups.** It is simple to verify that all the various filling-type properties described so far are preserved in moving from subgroup to supergroup. This immediately gives

Proposition 2.16. *If $H < G$ and $\{W_n\}$ is a filling or quasi-filling sequence in H then for any Følner sequence $\{F_n\}$ in G, $U^*_{(F,W)} \leq h$. If in addition $\frac{\log|W_{n+1}|}{\log|W_n|} \to 1$ the $T_*^{(F,W)} = T^*_{(F,W)} = h$.*

To obtain the same result for U_*, more care is needed, because one cannot automatically transfer an interpolation sequence from a subgroup to supergroup; this is due to the fact that condition (3) of the definition of interpolation sequences (definition 1.10) depends on the Følner sequence involved, and not only on the window sequence, and the Følner sequences in the supergroup are quite different from those in the subgroup. However, if $\{Y_n\}$ satisfies (1) and (2) of the definition of interpolation sequences with respect some to $\{W_n\}$ and in addition $\lim_{n\to\infty} \frac{\log|Y_n|}{\log|Y_{n-1}|} = 1$, then (3) of the definition of interpolation sequences is satisfied for *any* sequence $\{F_n\}$, and so $\{Y_n\}$ is an interpolation sequence in G for $\{F_n\}, \{W_n\}$ for *any* sequence $\{F_n\}$ in G. This gives

Proposition 2.17. *If $H < G$ and $\{W_n\}$ is an increasing incompressible sequence in H, and there exists a sequence $\{Y_n\}$ satisfying (1) and (2) of definition 1.10 and such that $\frac{\log|W_{n+1}|}{\log|W_n|} \to 1$, then for any Følner sequence $\{F_n\}$ in G we have $U^{(F,W)*} = U_*^{(F,W)} = h$.*

Corollary 2.18. *If G contains an element of infinite order then there exist in G window sequences $\{W_n\}$ for which for any tempered Følner sequence $\{F_n\}$, $U_*^{(F,W)} = U^{(F,W)*} = T_*^{(F,W)} = T^{(F,W)*} = h$.*

Proof. If $\langle g\rangle \cong \mathbb{Z}$ take $W_n = [-n; n] \subseteq \langle g\rangle$; since $\{W_n\}$ is incompressible in $\mathbb{Z}$ the same is true in G, and it is its own interpolation sequence. $\square$

We remark that in the case that $\{W_n\}$ is contained in some proper subgroup H of G, although we are looking at recurrence in H, the patterns we are looking at come from a Følner sequence in G, which is not contained in H. We emphasize that these results do not follow from an application of our recurrence theorems to the H-process which arises from the restriction of the G-action to H; in fact the H-process derived from an ergodic G-process need not be ergodic, and even if it is it need not have the same entropy as the original process. This observation means that we cannot weaken the requirement that F_n be a Følner sequence. If for example $\{F_n\}$ were a Følner sequence in some proper subgroup H of G, our results as applied to the H-process and the ergodic component to which x belongs would show that U_*, U^* etc. do converge, but their value depends on x and may be different from the entropy of the original process.

2.7. **Return times and densities may misbehave for "nice" window sequences which aren't filling.** It is natural to wonder which of the theorems about return times and densities are true under weaker conditions than those stated, and in particular whether the requirement that $\{W_n\}$ be filling may not be replaced with a weaker condition. One condition which has been used successfully in the proof of other pointwise theorems is that of being a Templeman sequence, ie an increasing Følner sequence with $|W_n^{-1}W_n| \leq C|W_n|$ (note that this is implies temperedness of the sequence $\{W_n\}$). In this section we show that even a stronger condition is not enough, and that the almost-certain bound

$$\liminf_k \inf_n -\frac{1}{|F_k|} \log U_{k,n}^{(F,F)}(x) \leq h$$

may fail even if $\{F_n\}$ is a symmetric, increasing Templeman sequence of all orders, ie for all d there is a constant C_d such that $|(F_n)^d| \leq C_d|F_n|$ for all n.

Let $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in the complex plane and let $Tz = e^{2\pi i\theta}z$ be an irrational rotation, ie $\theta \in \mathbb{R} \setminus \mathbb{Q}$. Let $X : S^1 \to \{0,1\}$ be the map $z \mapsto \operatorname{sgn} \operatorname{Im} z$, and consider the process $X_n = T^n X$ defined on S^1 when S^1 is equipped with Lebesgue measure. It is well known that this process is ergodic and has zero entropy. Our goal here is to construct a sequence of increasing, symmetric finite sets $F_n \subseteq \mathbb{Z}$ such that $\{F_n\}$ is Templeman of all orders and $|F_n| = 2n$, but for which there is a sequence of indices k_i such that $T_{k_i}^{(F,F)}(z) > 2^{|F_{k_i}|}$ for every $z \in S^1$, so $\frac{1}{|F_k|} \log T_k^{(F,F)}$ does not converge even in probability to the entropy. This implies that the asymptotic upper bound for $\inf_n -\frac{1}{|F_k|} \log U_{k,n}^{(F,F)}$ fails as well.

Let $\varepsilon_n \searrow 0$ be a decreasing sequence of positive numbers such that for every $z \in S^1$, if $0 < |k| < n$ then $|T^k z - z| > \varepsilon_n$. For any $z_0 \in S^1$ one can find such an ε_n, since T has no periodic orbits. Since T is an isometry of S^1 such a choice of ε_n is good for all $z \in S^1$ if it is good for z_0.

Let ℓ_n be an increasing sequence of positive integers such that any realization $(X_{-\ell_n}(z), \ldots, X_{\ell_n}(z))$ determines z up to a distance of ε_n. More precisely, we require that if $z, z' \in S^1$ satisfy $X_k(z) = X_k(z')$ for all $-n \leq k \leq n$ then $|z - z'| < \varepsilon_n$. Such ℓ_n exist because if we denote by $\mathcal{P}$ the partition of S^1 determined by X then $\vee_{-\ell}^{\ell} T^n \mathcal{P}$ partitions S^1 into segments whose length tends to 0 as $\ell \to 0$.

Finally, write

$$J_n = \left\{ k \in \mathbb{Z} : |z - T^k z| > \varepsilon_n \text{ (for any or all } z \in S^1) \right\}.$$

Note that J_n are symmetric (because T is an isometry) and that since ε_n are decreasing, $J_n \subseteq J_{n+1}$. Also note that if $I = [a; b] \subseteq \mathbb{Z}$ then

$$k, m \in I \setminus J_n \Rightarrow |z - T^{k-m} z| = |T^k z - T^m z| < 2\varepsilon_n$$

and so if $\varepsilon_{n'} > 2\varepsilon_n$ then $n' < |k - m|$. Thus if $|I| < n'$ then $|J_n \cap I| \geq |I| - 1$. It follows that the relative density of J_n in any segment I of length $n' - 1$ is at least $1 - \frac{1}{n'-1}$.

We define the sets F_n inductively. At each stage we obtain F_n from F_{n-1} by appending to F_{n-1} some symmetric pair of numbers $\pm m$. Given F_n and a symmetric set $E \subseteq \mathbb{Z}$, when we will say that we "add E to F_n" we will mean that over several stages we add to F_i the smallest pair of numbers (in absolute value) in $E \setminus F_i$ until E is exhausted or some condition is met.

We start with $F_0 = \emptyset$ and add to it the segment $[-\ell_1; \ell_1]$ to obtain the sequence $F_1, \ldots, F_{k_1}$. We now add J_1 until we obtain a set F_{n_1} such that $|F_{n_1}| > 2^{|F_{k_1}|}$. Now we add $[-\ell_2; \ell_2]$ to F_{n_1} and obtain F_{k_2}; then add J_2 to F_{k_2} until we obtain a set F_{n_2} with $|F_{n_2}| > 2^{|F_{k_2}|}$. We proceed in this manner, alternately adding $[\ell_{-i}; \ell_i]$ to obtain F_{k_i} and then J_i to obtain F_{n_i} with $|F_{n_i}| > 2^{|F_{k_i}|}$.

The sequence $\{F_n\}$ thus constructed is clearly increasing and symmetric, and satisfies $|F_n| = 2n$. Next, we claim that $T_{k_i}(z) > 2^{|F_{k_i}|}$ for every $z \in S^1$. To see that this is true for all i, note that $z(F_{k_i})$ determines z up to ε_i because $[-\ell_i; \ell_i] \subseteq F_{k_i}$, while for every $m \in F_{n_i} \setminus F_{k_i}$ we have $m \in J_i$, so $|z - T^m z| > \varepsilon_i$. Therefore $T^m z([-\ell_i; \ell_i]) \neq z([-\ell_i; \ell_i])$ and this forces $T^m z(F_{k_i}) \neq z(F_{k_i})$. Since $|F_{n_i}| > 2^{|F_{k_i}|}$, this proves the claim.

As for $\{F_n\}$ being Templeman, it suffices to show that $\lim_{n\to\infty} \frac{|F_n|}{\max F_n - \min F_n} = 1$; this implies both that $\{F_n\}$ is a Følner sequence, and that $|F_n^d| \leq C_d |F_n|$ for some constant C_d, for all n. But this follows from our remarks about the relative density of J_n in certain segments.

We remark that it seems to be more difficult to construct an example of such a sequence $\{F_n\}$ for which the lower bound for $U_{k,n}$ fails. Perhaps the lower bound $U_* \geq h$ holds more generally than the upper bound, as happens in the case of T_k. we have so far been unable to determine whether this is so.

3. Upper bounds for return times and densities

In this section we prove the upper bounds $T^* \leq h$ and $U^* \leq h$ as described in the introduction.

3.1. The upper bound for U^*.

We begin with some further remarks about the definition of $U_{k,n}$. Recall that in section 1.4 we defined $U_{k,n}$ to be size of the maximal collection of repetitions of $x(F_k)$ in $x(W_n)$ satisfying (a) the collection includes the trivial repetition $x(F_k)$ itself, and (b) the members of the collection are pairwise disjoint. If we drop these restriction we obtain the quantity

$$V_{k,n}(x) = V_{k,n}^{(F,W)}(x) = \frac{1}{|W_n|} \#\{f \in W_n \,:\, x(F_k) \text{ repeats at } f\}$$

which is related to $U_{k,n}$ by

$$U_{k,n} \leq V_{k,n} \leq |F_k^{-1} F_k| U_{k,n} \leq |F_k|^2 U_{k,n}. \tag{3.1}$$

The middle inequality is due to the fact that from any collection of right translates of F_k we can obtain a pairwise disjoint sub-collection at least $\frac{1}{|F_k^{-1}F_k|}$ the size of the original collection which includes any single prescribed member of the original collection (which in our case we take to be

the set F_k, representing the original pattern). From this relation we see that in order to prove that $U^* \leq h$, it suffices to prove that $V^* = \limsup_n \sup_k -\frac{1}{|F_k|}\log V_{k,n} \leq h$.

The key lemma which makes filling sequences special is the following, which states that if a set E has a low "local" density with respect to $\{W_n\}$ then it has a low "global" density:

Lemma 3.1. *Suppose $\{W_n\}$ is a C-filling sequence and $E \subseteq G$ finite. If for each $f \in E$ there is an index $n(f)$ such that $\frac{1}{|W_{n(f)}|}|W_{n(f)}f \cap E| \leq \alpha$ then*

$$\frac{|E|}{|\bigcup_{f\in E} W_{n(f)}f|} \leq \frac{\alpha}{C}.$$

Note that we always have $0 < C \leq 1$.

Proof. Order E by decreasing value of $n(\cdot)$: $E = \{f_1,\ldots,f_I\}$ and write $n(f_i) = n(i)$. We select a subsequence $\left\{f_{i(j)}\right\} \subseteq \{f_1,\ldots,f_I\}$ inductively. Set $i(1) = 1$ and let $i(j)$ be the first index such that $f_{i(j)} \notin \bigcup_{k<j} W_{n(k)}f_k$; the process ends after J steps when $E \subseteq \bigcup_{j=1}^{J} W_{n(i(j))}f_{i(j)}$ (recall that $1_G \in W_n$ for each n, since $\{W_n\}$ is filling). Clearly $\left\{W_{n(i(j))}f_{i(j)}\right\}_{j=1}^{J}$ is incremental. Now

$$|E| \leq \sum_{j=1}^{J} |W_{n(i(j))}f_{i(j)} \cap E| \leq \alpha \sum_{j=1}^{J} |W_{n(i(j))}| \leq \frac{\alpha}{C}|\bigcup_{j=1}^{J} W_{n(i(j))}f_{i(j)}|$$

(the middle inequality is because by hypothesis $|W_{n(i)}f_i \cap E| \leq \alpha|W_{n(i)}|$ for every $i = 1,\ldots,I$ and the last because $\{W_n\}$ is C-filling). Since

$$\bigcup_{j=1}^{J} W_{n(i(j))}f_{i(j)} \subseteq \bigcup_{i=1}^{I} W_{n(i)}f_i$$

the lemma follows. $\square$

We apply lemma 3.1 in the following situation: Consider an ergodic G-process defined on $(\Omega,\mathcal{F},P)$ with entropy h. Suppose $x \in \Omega$, $k < L << N$ are fixed, and $\varphi : F_k \to \Sigma$ is some F_k pattern. Let E be the set of repetitions of φ in $x(F_N)$ with the additional property that near them φ repeats not frequently enough with respect to $V_{k,n}$:

$$E = \left\{ f \in F_N \;\middle|\; \begin{array}{c} fx \in [\varphi] \text{ and there exists } n(f) \leq L \\ \text{s.t. } V_{k,n(f)}(fx) < 2^{-(h+\varepsilon)|F_k|} \end{array} \right\}.$$

Since for any $f \in E$ we have $fx \in [\varphi]$ and $\frac{1}{|W_{n(f)}|}|E \cap W_{n(f)}f| < 2^{-(h+\varepsilon)|F_k|}$, from the lemma we see that

$$\frac{1}{|\cup_{f\in E} W_{n(f)}f|}|E| < \frac{1}{C} \cdot 2^{-(h+\varepsilon)|F_k|}.$$

If we assume that N is large enough so that $|(\cup_{n\leq L}W_n)F_N| \leq 2|F_N|$ (this happens eventually since $\{F_n\}$ is a Følner sequence), then since $\cup_{f\in E}W_{n(f)}f \subseteq (\cup_{n\leq L}W_n)F_N$, we see that

$$\frac{1}{|F_N|}|E| \leq \frac{2}{|\cup_{f\in E} W_{n(f)}f|}|E| < \frac{2}{C}\cdot 2^{-(h+\varepsilon)|F_k|}.$$

Given some $\Phi \subseteq \Sigma^{F_k}$ and repeating this argument for every $\varphi \in \Phi$ we have

$$\frac{1}{|F_N|}\#\left\{ f\in F_N \;\middle|\; \begin{array}{c} fx\in[\Phi] \text{ and there exists } n(f)\leq L \\ \text{s.t. } V_{k,n(f)}(fx) < 2^{-(h+\varepsilon)|F_k|} \end{array} \right\} < \frac{2}{C}\cdot 2^{-(h+\varepsilon)|F_k|}\cdot|\Phi|.$$

Theorem 3.2. *If $\{F_n\}$ is a tempered Følner sequence and if $\{W_n\}$ is filling, then*

$$U^*(x) = \limsup_{k\to\infty}\sup_n -\frac{1}{|F_k|}\log U_{k,n}^{(F,W)}(x) \leq h$$

almost surely.

Proof. It is enough to prove that

$$V^* = \limsup_k \sup_n -\frac{1}{|F_k|}\log|V_{k,n}| \leq h$$

almost surely. Suppose to the contrary that for some $\varepsilon > 0$ and measurable non-null set $B \subseteq \Omega$, for every $x \in B$

$$\limsup_{k\to\infty}\sup_n -\frac{1}{|F_k|}\log V_{k,n}^{(F,W)}(x) > h+\varepsilon.$$

Let $P(B) > p > 0$. Then from the SMB theorem, there is a measurable subset $B_0 \subseteq B$, $P(B_0) > p$, a sequence $\{\Phi_k\}$ of sets $\Phi_k \subseteq \Sigma^{F_k}$ and integer L_1 such that for $k > L_1$ it holds that $|\Phi_k| \leq 2^{(h+\varepsilon/2)|F_k|}$, and $x(F_k) \in \Phi_k$ for any $x \in B_0$. It may further be assumed that for some $L_2 >> L_1$, for any $x \in B_0$ there are indices $k(x), n(x)$ with $L_1 \leq k(x) \leq L_2$ and $n(x) \leq L_2$ such that $V_{k(x),n(x)}(x) < 2^{-(h+\varepsilon)|F_{k(x)}|}$.

Choose N very large with respect to L_2 so that F_N is very W_{L_2}-invariant, and large enough so that there is an $x \in \Omega$ such that

$$\frac{1}{|F_N|}\#\{f\in F_N \,:\, fx\in B_0\} > p$$

(by the ergodic theorem such x's are bound to exist for N large enough).

Fix $x \in \Omega$ for which the above holds. For $f \in F_N$ such that $fx \in B_0$, write $k(f) = k(fx)$, $n(f) = n(fx)$. Define

$$E = \{f\in F_N \,:\, fx\in B_0\},$$
$$E_k = \{f\in F_N \,:\, fx\in B_0 \text{ and } k(f)=k\}.$$

Clearly $E \subseteq (\cup_{k=L_1}^{L_2} E_k)$, so by our choice of x we should have

$$\frac{1}{|F_N|}|\cup E_k| \geq p.$$

On the other hand, according to the discussion after lemma 3.1,

$$\frac{1}{|F_N|}|E_k| \leq \frac{2}{C}2^{-(h+\varepsilon)|F_k|}|\Phi_k| \leq \frac{2}{C}(2^{-\varepsilon/2})^{|F_k|}$$

so that

$$\frac{1}{|F_N|}|\cup E_k| = \frac{1}{|F_N|}\sum_{L_1<k<L_2}|E_k| \leq C' \cdot (2^{-\varepsilon/2})^{F_{L_1}} \tag{3.2}$$

for some constant C' depending only on ε, so if L_1 was chosen large enough with respect to ε (recall our assumption that that $|F_n| \geq n$), this is less than p, a contradiction. □

Esssencially the same proof gives us

Theorem 3.3. *Let $\{F_n\}$ be a tempered Følner sequence in G. If $\{W_n\}$ is a sequence such that W_n is C_n-filling and $\{n(k)\}$ is a sequence of integers such that $\sum_{k=1}^{\infty} \frac{2^{-\alpha|F_k|}}{C_{n(k)}} < \infty$ for every $\alpha > 0$, then for almost every x,*

$$\limsup_{k\to\infty} -\frac{1}{|F_k|}\log U_{k,n(k)}^{(F,W)}(x) \leq h.$$

The proof uses the fact that lemma 3.1 has an analogue when $\{W_n\}$ is replaced with a single C-filling set W, namely that if for a finite subset $E \subseteq G$ we have $\frac{|Wf\cap E|}{|W|} < \alpha$ for every $f \in E$ then $\frac{|E|}{|WE|} \leq \frac{\alpha}{C}$ (to see this, apply the lemma with $W_n = W$). Using this the proof now follows exactly the lines of the proof of theorem 3.2; the relation $\sum \frac{2^{-\alpha|F_k|}}{C_{n(k)}} < \infty$ is used to show that the middle term of equation (3.2) is small when L_1 is large enough.

3.2. **The upper bound for T^*.** In order for $T_k^{(F,W)}$ to be meaningful the window sequence $\{W_n\}$ must grow slowly enough to detect a too-soon return. What slowly means here is that for every h there exist window sets of size approximately $2^{h|F_k|}$ for large enough k. One condition which ensures this without reference to the sequence $\{F_n\}$ is that $\frac{\log|W_{n+1}|}{\log|W_n|} \to 1$. With this condition in place, one can deduce the bound for T^* from the bound for U^*:

Lemma 3.4. *Suppose $\{W_n\}$ is an increasing sequence such that $\frac{\log|W_{n+1}|}{\log|W_n|} \to 1$. For $x \in \Omega$, if $U^*(x) \leq h$ then $T^*(x) \leq h$.*

Proof. If $T_k(x) = |W_{R_k(x)}| > 2^{(h+\varepsilon)|F_k|}$ for infinitely many k, then the growth condition implies that $|W_{R_k(x)-1}| \geq 2^{(h+\varepsilon/2)|F_k|}$ for infinitely many k. But for such a k the only occurence of $x(F_k)$ in

$W_{R_k(x)-1}$ is the original pattern, so $|U_{k,R_k(x)-1}(x)| \leq 2^{-(h+\varepsilon/2)|F_k|}$ for infinitely many k, which is impossible because $U^*(x) \leq h$. $\square$

Theorem 3.5. *Let $\{W_n\}$ be a quasi-filling sequence satisfying $\frac{\log|W_{n+1}|}{\log|W_n|} \to 1$. Then for any tempered Følner sequence $\{F_n\}$, $T^*_{(F,W)} \leq h$ almost surely.*

Proof. Let $\varepsilon > 0$. Define

$$n(k) = \min\left\{n : |W_n| > 2^{(h+\varepsilon)|F_k|}\right\}.$$

For $\alpha > 0$ we have, since $\{W_n\}$ is quasi-filling, that

$$\sum \frac{2^{-\alpha|F_k|}}{C_{n(k)}} \leq \sum(\frac{1}{C_{n(k)}}2^{(h+\varepsilon)|F_k|})^{-\alpha/(h+\varepsilon)} \leq \sum \frac{|W_{n(k)}|^{-\alpha/(h+\varepsilon)}}{C_{n(k)}} < \infty$$

and this now gives, via theorem 3.3, that

$$\limsup_k \frac{1}{|F_k|} \log U^{(F,W)}_{k,n(k)}(x) \leq h$$

almost surely. This in turn implies that $T^* \leq h$ almost surely, by reasoning like that of the lemma 3.4 above. $\square$

4. Lower bounds on return times and densities

In this section we prove the bounds $T_* \geq h$ and $U_* \geq h$. The bound $T_* \geq h$ is proved by constructing a code for the process which beats entropy if $T_* \not\geq h$. This is close in spirit to the proof of $T_* \geq h$ originally given by Ornstein and Weiss for the cases $G = \mathbb{Z}$ and $G = \mathbb{Z}^d$, but some new combinatorial machinery is needed to make the construction possible for general groups and general window sets. We first give a brief description of the coding ideas we will need, and then explain the combinatorics and detailed construction of the code. We conclude with the proof of the bound $U_* \geq h$, which relies both on the filling ideas used in the last section and on the lower bound $T_* \geq h$.

4.1. **Coding and entropy.** For completeness we take a detour and discuss the connection between entropy and efficient coding of a process.

A useful characterization of the entropy of a process is as the lower bound of all achievable coding rates for the process. An F_n-*code* is a map $c : \Sigma^{F_n} \to \{0,1\}^* = \cup_{n\geq 0}\{0,1\}^n$. Thus c encodes each pattern $\varphi : F_n \to \Sigma$ in a binary string (=sequence) $c(\varphi)$, which is called the *codeword* associated with φ. If c is an injection the coding is said to be *invertible or faithful.* For $x \in \Omega$ and an F_n-code c, we write $c(x)$ instead of $c(x(F_n))$. If c_n is an F_n-code, we say that $\{c_n\}$ is an $\{F_n\}$-code. $\{c_n\}$ is invertsble if every c_n is.

The number of bits (=symbols) in the codeword $c(\varphi)$ of φ is called the *length* of the codeword $c(\varphi)$ and is denoted by $\ell(c(\varphi))$. Note that the codeword length of different patterns may vary. For an F_n-code c and process $\{X_g\}$, the average number of *bits per symbol* used by the code is the quantity $\rho(c) = \frac{1}{|F_n|}\int \ell(c(x))dP(x)$; this is the average number of bits used in the c-encoding of a random pattern $x(F_n)$, normalized by the number of symbols in the F_n-pattern $x(F_n)$.

For $r \geq 0$, we say that r is an *achievable coding rate* for the process $\{X_g\}$ if there exists an invertable $\{F_n\}$-code $\{c_n\}$ such that $\limsup \rho(c_n) \leq r$. This means we can represent arbitrarily large patterns from the process using approximately r bits for every symbol of the pattern. The *coding rate* of the process $\{X_g\}$ is the infimum of achievable coding rates:

$$R = \inf\left\{r \geq 0 \,:\, r \text{ is an achievable coding rate for the process } \{X_g\}\right\}.$$

It is not hard to see that the coding rate of a process is the same as the process entropy h (incidentally, this shows that R does not depend on the sequence $\{F_n\}$). To see that $R \geq h$, let $\{c_n\}$ be a sequence of invertable codes such that $\limsup_{n\to\infty} \rho(c_n) < r$. Then there is some $\varepsilon > 0$ such that for infinitely many n there is a measurable set $A_n \subseteq \Omega$ with $P(A_n) \geq \varepsilon$, and $\frac{1}{|F_n|}\ell(c_n(x)) < r$ for every $x \in A_n$. Since c_n is an injection, and for every $x \in A_n$ we have $\ell(c_n(x)) \leq r|F_n|$, we have

$$\{c_n(x) \,:\, x \in A_n\} \subseteq \bigcup_{k<r|F_n|} \{0,1\}^k$$

so every $x \in A_n$ belongs to at most one $\sum_{k\leq r|F_n|} 2^k \leq 2 \cdot 2^{r|F_n|}$ F_n-atoms. The Shannon-McMillan theorem now implies that $h \leq r$.

To see that $h \geq R$ we formulate the following simple lemma, which enables us to turn a non-faithful code sequence into a faithful code sequence.

Lemma 4.1. *Suppose that $\{c_n\}$ is a sequence of F_n-codes and $\Phi_n \subseteq \Sigma^{F_n}$ is such that $c_n|_{\Phi_n}$ is an injection and $\frac{1}{|F_n|}\ell(c_n(\varphi)) \leq r$ for every $\varphi \in \Phi_n$. Assume also that $P([\Phi_n]) \to 1$. Then $R \leq r$.*

Proof. Let e_n be an enumeration coding of Σ^{F_n}, ie it assigns to $\varphi \in \Sigma^{F_n}$ a binary string representing the index of φ in some fixed enumeration of Σ^{F_n}. The code e_n is faithful. Now define a code $\{\hat{c}_n\}$ by

$$\hat{c}_n(\varphi) = \begin{cases} 0c_n(\varphi) & \varphi \in \Phi_n \\ 1e_n(\varphi) & \varphi \notin \Phi_n \end{cases}$$

$\{\hat{c}_n\}$ is clearly an invertible code. Moreover,

$$\frac{1}{|F_n|}\ell(\hat{c}_n(\varphi)) = \begin{cases} \frac{1}{|F_n|} + \frac{1}{|F_n|}\ell(c_n(\varphi)) & \varphi \in \Phi_n \\ \frac{1}{|F_n|} + \lceil \log |\Sigma| \rceil & \varphi \notin \Phi_n \end{cases}$$

so,

$$\begin{aligned}\rho(\widehat{c}_n) &= \frac{1}{|F_n|}\int \ell(\hat{c}_n(x))dP \\ &\leq \frac{1}{|F_n|} + r\cdot P([\Phi_n]) + \lceil \log|\Sigma|\rceil\, P(\Omega\setminus[\Phi_n]) \\ &\to r\end{aligned}$$

as desired. □

If we choose $\Phi_n \subseteq \Sigma^{F_n}$ in such a way that $\frac{1}{|F_n|}\log|\Phi_n| \to h$ and $P([\Phi_n]) \to 1$, choose c_n to be an enumeration coding of Φ_n and define it arbitrarily of $\Sigma^{F_n}\setminus\Phi_n$, then the lemma implies that $\rho(\widehat{c}_n) \leq h+\varepsilon$ for every $\varepsilon > 0$, so $R \leq h$.

We conclude this section with a combinatorial lemma which will be used later. Its importance is that it enables us to describe small random subsets of F_n at a low rate by using an enumeration code.

Lemma 4.2. *Let $0 < \lambda < 1/2$ and F a finite set. The number of subsets of F whose size is at most $\lambda|F|$ is bounded by $2^{\delta(\lambda)|F|}$, where* $\lim_{\lambda\to 0}\delta(\lambda) = 0$.

Proof. The number of such subsets is

$$\sum_{m<\lambda|F_n|}\binom{|F_n|}{m} \leq 2^{(-\lambda\log\lambda-(1-\lambda)\log(1-\lambda))|F|}.$$

The bound follows from Stirling's formula; or see [9, p. 52] for an elementary proof. □

4.2. **The lower bound for T_* (Part I).** In this section we begin the proof of

Theorem 4.3. *If $\{F_n\}$ is a tempered sequence and $\{W_n\}$ is an increasing sequence of finite subsets of G, then*

$$T_*^{(F,W)} = \liminf_{k\to\infty}\frac{1}{|F_k|}\log T_k^{(F,W)} \geq h$$

almost surely.

We will need some notation.

Definition 4.4. Let $\{F_n\}$ be a fixed sequence. A *cover* ν is a collection of sets of the form F_nf for some $n\in\mathbb{N}$ and $f\in G$. The set $E\subseteq G$ of f's such that some F_nf is represented in ν is called the *set of centers of* ν and is denoted by $\operatorname{dom}\nu$. We also say then that ν is a cover *over* E.

It would be more precise to say that a cover ν over E is a collection of pairs $(n,f)\in\mathbb{N}\times E$, because we wish to allow more than one set F_nf over each $f\in E$, and also since the set F_nf does

not, in general, determine n and f uniquely. However, it is more convenient to think of ν as a (multi)set with elements of the form $F_n f$, and this is what we will do.

We write $\cup\nu$ for the union of members of ν. When we wish to specify the elements of ν by name, we will write $\cup_\nu F_n f$, $\sum_\nu |F_n f|$, etc. This is imprecise because the index n depends on f but we omit this when convenient.

Fix $\{F_n\}$ and $\{W_n\}$. Let $x \in \Omega$, $f \in G$ and $n \in \mathbb{N}$, and suppose $R_n(fx) < \infty$. There then exists an element $f' \in G$ at which $fx(F_n)$ repeats, and is closest to f with respect to $\{W_n\}$. To be precise, there is an f' such that

(1) $f'x \in [fx(F_n)]$ (ie the F_n-pattern at f in x repeats in f').
(2) $f'f^{-1} \in W_{R_n(fx)}$ (ie the offset from f to f' is in $W_{R_n(f(x))}$).

In general there may be more than one f' with these properties. For fixed x and n, we assume some choice has been made, and denote this element by $f^{x,n}$.

For a set $F_n f$ such that $R_n(fx) < \infty$, we write $(F_n f)^x = F_n f^{n,x}$. The (partial) map $F_n f \mapsto (F_n f)^x$ is called the *displacement arising from x*. Strictly speaking, the map $F_n f \mapsto (F_n f)^x$ is not well defined since the set $F_n f$ does not uniquely determine n and f. However, n and f will always be clear from the context so no confusion should arise. Finally, for a cover ν such that $R_n(fx) < \infty$ for every $F_n f \in \nu$ we write $\nu^x = \{(F_n f)^x : F_n f \in \nu\}$. Thus ν^x is a cover over the set $\{f^{x,n} : F_n f \in \nu\}$. One should think of the displacement ν^x of ν as the cover obtained from ν by moving each set $F_n f \in \nu$ to one of the nearest repetitions of the pattern induced on it by x.

We will only be working with covers ν and $x \in \Omega$ such that $R_n(fx) < \infty$ for every $F_n f \in \nu$, and avoid further mention of this fact.

We can now state the idea behind the proof of theorem 4.3. Suppose by way of contradiction that $P(T_* < h - \varepsilon) > p > 0$. We will try to construct an F_N-code c_N which beats entropy. By the ergodic theorem, for almost every x and large N at least a p-fraction of the $f \in F_N$ are such that $T_n(fx) \leq 2^{(h-\varepsilon)|F_n|}$ for some $n = n(f) \in \mathbb{N}$. In order to describe the pattern $x(F_n f)$ it is enough to describe the pattern $x(F_n f^{x,n})$ and the element $f^{x,n} f^{-1}$. Since the latter is in $W_{R_n(fx)}$ and this set is of size at most $2^{(h-\varepsilon)|F_n|}$, we can describe it in approximately $(h-\varepsilon)|F_n|$ bits. Thus if the pattern $x(F_n f^{x,n})$ were known, the pattern $x(F_n f)$ could be coded in $h - \varepsilon$ bits per symbol.

Let $E \subseteq F_N$ be the set of $f \in F_N$ for which there exists such an $n(f)$ with $R_{n(f)}(fx) \leq h - \varepsilon$. Let $\nu = \left\{F_{n(f)} f\right\}_{f \in E}$. If it should happen that (a) the collection ν is pairwise disjoint, (b) $\cup\nu$ is disjoint from $\cup\nu^x$, (c) we can code $x(F_N \setminus \cup\nu)$ at, say $h + p\varepsilon/2$ bits per symbol, and (d) $\cup\nu$ is at least a p-fraction of F_N, then the discussion above shows that we can code $x(\cup\nu)$ at around $h - \varepsilon$ bits per symbol. Given that the above hold, we can then code $x(F_N)$ at a rate of $h - p\varepsilon/2$ bits per symbol.

Conditions (c) and (d) are not to difficult: (d) would follow from (a) and the fact that E is a p-fraction of F_N, and (c) can be achieved because $F_N \setminus \cup\nu$ can be mostly covered by entropy-typical sets when N is large, and a standard argument then gives that $x(F_N \setminus \cup\nu)$ can be coded at a little more than h bits per symbol. The difficulty in realizing this outline is that in (a) and (b) cannot in general be achieved.

The proof of the return times theorem for $\mathbb{Z}$ and $\mathbb{Z}^d$ given in [8] avoids this problem by allowing the elements of ν and ν^x to intersect arbitrarily, and employs a more elaborate coding scheme which uses the order structure of $\mathbb{Z}^d$ in an essential way. This method doesn't generalize to other groups.

For our proof we will drop the disjointness requirements (a) and (b) and instead require only "almost disjointness". To be precise,

Definition 4.5. Let $\varepsilon > 0$. A sequence $\{H_i\}_{i\in I}$ of finite subsets of G are said to be ε-*disjoint* if there exist subsets $H_i' \subseteq H_i$ such that $\{H_i'\}_{i\in I}$ are pairwise disjoint and $|H_i'| \geq (1-\varepsilon)|H_i|$. We say that a cover ν is ε-disjoint if the collection $\{F_n f \in \nu\}$ is ε-disjoint.

If $\{H_i\}_{i=1}^I$ has the property that for each $i \leq I$, $|H_i \cap (\cup_{j<i} H_j)| \leq \varepsilon |H_i|$ then the collection is ε-disjoint; simply set $H_i' = H_i \setminus \cup_{j<i} H_j$.

An important property of ε-disjoint collections is that the size of their union is almost the sum of their sizes. To be precise, let $\{H_i\}_{i\in I}$ be ε-disjoint and $H_i' \subseteq H_i$ as in the definition. Then we have

$$|\bigcup_{i\in I} H_i| \geq |\bigcup_{i\in I} H_i'| = \sum_{i\in I} |H_i'| \geq (1-\varepsilon)\sum_{i\in I}|H_i|.$$

Clearly for two ε-disjoint collections $\{D_i\}$ and $\{E_j\}$, if $D_i \cap E_j = \emptyset$ for every i, j then $\{D_i\} \cup \{E_j\}$ is an ε-disjoint collection; and so for any funion of pairwise disjoint ε-disjoint collections.

Definition 4.6. For $x \in \Omega$ and a cover $\nu = \left\{F_{n(i)} f_i\right\}_{i=1}^I$ over $E = \{f_1, \ldots, f_I\}$, we say that the pair (ν, ν^x) are ε-disjoint if for every $1 \leq i \leq I$,

$$\left| F_{n(i)} f_i \cap \bigcup_{j<i} \left(F_{n(j)} f_j \cup \left(F_{n(j)} f_j\right)^x\right)\right| \leq \varepsilon |F_{n(i)} f_i|.$$

According to the discussion before the definition, if (ν, ν^x) is ε-disjoint then ν is ε-disjoint. Note that in order for the statement that (ν, ν^x) is ε-disjoint to be meaningful, the set of centers of ν must be linearly ordered.

The point of this definitions is the following:

Lemma 4.7. *Let $0 < \varepsilon < \frac{1}{2}$, $x \in \Omega$ and suppose $\nu = \left\{F_{n(i)} f_i\right\}_{i=1}^I$ is a cover over $E = \{f_1, \ldots, f_I\}$ such that (ν, ν^x) is ε-disjoint. Then there exists a set $A \subseteq \cup\nu$ satisfying $|A| < 2\varepsilon |\cup \nu|$ and such*

that, given the patterns $x(\cup\nu^x \setminus \cup\nu)$ and $x(A)$, and given the map $f_i \mapsto f^{x,n(i)}$, we can reconstruct the pattern $x(\cup\nu)$.

Proof. Set
$$H_i = F_{n(i)}f_i \setminus \bigcup_{j<i}(F_{n(j)}f_j)^x.$$
Since (ν,ν^x) is ε-disjoint, $|H_i| \geq (1-\varepsilon)|F_{n(i)}|$, and for $j > i$ we have $(F_{n(i)}f_i)^x \cap H_j = \emptyset$. Define
$$A = (\cup\nu) \setminus \bigcup_{i=1}^{I} H_i.$$
Using the fact that ν is ε-disjoint (since (ν,ν^x) is) ,
$$|A| < \varepsilon \sum |F_{n(i)}f_i| \leq \frac{\varepsilon}{1-\varepsilon}| \cup \nu| \leq 2\varepsilon | \cup \nu|.$$
It remains to show that given $x(A)$, $x(\cup\nu^x \setminus \cup\nu)$ and the map $f \mapsto f^{x,n(f)}$ we can deduce $x(\cup\nu)$.

We claim that for any i if we know $x(\cup_{j<i}F_{n(j)}f_j)$ then we can deduce $x(F_{n(i)}f_i)$; this will complete the proof. Since $f_i^{x,n(i)}x(F_{n(i)}) = f_i x(F_{n(i)})$, and we know f_i and $f_i^{x,n(i)}$, it suffices to show that we can deduce $x((F_{n(i)}f_i)^x)$. Now since by assumption we already know $x(\cup_{j<i}F_{n(j)}f_j)$, and since $F_{n(i)}f_i \cap (F_{n(i)}f_i)^x = \emptyset$, all we need do is find $x((F_{n(i)}f_i)^x \cap \bigcup_{j>i} F_{n(j)}f_j)$. But this is known, since $(F_{n(i)}f_i)^x \cap \bigcup_{j>i} F_{n(j)}f_j \subseteq A$, and we know $x(A)$. $\square$

In order to use the lemma 4.7 to produce good codes we need to know how to produce large covers ν such that (ν,ν^x) is ε-disjoint. We address this next.

4.3. **Disjointification lemmas.** Our methods in this section follow those in [11].

We first need some more notation. If ν,μ are covers, their join is
$$\nu \vee \mu = \{F_n f \,:\, F_n f \in \nu \text{ or } F_n f \in \mu\}$$
(we use this notation instead of the more natural $\nu \cup \mu$ to avoid confusion with the set $\cup\nu$ defined above in 4.4). The join of several covers $\nu_1,\ldots,\nu_k$ is denoted by $\bigvee_{i=1}^k \nu_i$.

The *restriction* of a cover ν over E to a subset of $E' \subseteq E$ is denoted by $\nu|_{E'} = \{F_n f \in \nu \,:\, f \in E'\}$. We say that μ is a *subcover* of ν if $\mu \subseteq \nu$, ie if $F_n f \in \nu$ whenever $F_n f \in \mu$. Thus if ν is a cover over E and $E' \subseteq E$ we have $\nu|_{E'} \subseteq \nu$.

We also write $\min\nu = \min\{n \,:\, F_n f \in \nu\}$, and similarly $\max\nu$.

Definition 4.8. Let $\{F_n\}$ be a Følner sequence. The *blowup* of F_n is $F_n^+ = (\cup_{k<n}F_k^{-1})F_n$. For a cover ν, write $\cup^+\nu = \cup_\nu F_n^+ f$.

The point of this is that if $F_k f$ and $F_n g$ are two translates with $k < n$, then $f \notin F_n^+ g$ iff $F_k f \cap F_n g = \emptyset$. Thus if $F_k f$ is a set and ν some cover, and $k < \min\nu$, then $f \notin \cup^+\nu$ iff $F_k f \cap (\cup\nu) = \emptyset$.

For a Følner sequence $\{F_n\}$, the condition $|F_n^+| \leq C|F_n|$ is equivalent to saying that $\{F_n\}$ is tempered with constant C.

The following properties will be useful. They state that if $\{F_n\}$ is tempered then the size of $\cup^+\nu$, $\cup(\nu \vee \nu^x)$ and $\cup^+(\nu \vee \nu^x)$ are at most a constant times the size of $\cup\nu$, assuming some disjointness criteria from ν or (ν, ν^x).

Lemma 4.9. *Suppose $\{F_n\}$ is tempered with constant C. If ν is an ε-disjoint cover then $|\cup^+ \nu| \leq \frac{C}{1-\varepsilon}|\cup \nu|$.*

Proof. We have

$$|\cup^+ \nu| = |\bigcup_{\nu} F_n^+ f| \leq \sum_{\nu} |F_n^+ f| \leq C \sum_{\nu} |F_n f| \leq \frac{C}{1-\varepsilon}|\cup \nu|$$

(the last inequality follows from the fact that ν is ε-disjoint). □

Lemma 4.10. *If (ν, ν^x) is ε-disjoint then $|\cup(\nu \vee \nu^x)| \leq \frac{2}{1-\varepsilon}|\cup \nu|$. If in addition $\{F_n\}$ is tempered with constant C then $|\cup^+ (\nu \vee \nu^x)| \leq \frac{2C}{1-\varepsilon}|\cup \nu|$.*

Proof. If (ν, ν^x) is ε-disjoint then it is clear from the definitions that ν is ε-disjoint. We therefore have

$$|\cup(\nu \vee \nu^x)| \leq \sum_{\nu}(|F_n f| + |(F_n f)^x|) = 2\sum_{\nu} |F_n f| \leq \frac{2}{1-\varepsilon}|\cup \nu|.$$

If $\{F_n\}$ is tempered with constant C,

$$|\cup^+ (\nu \vee \nu^x)| \leq 2\sum_{\nu} |F_n^+ f| \leq 2C\sum_{\nu} |F_n f| \leq \frac{2C}{1-\varepsilon}|\cup \nu|.$$

□

A cover ν over E is constant if for some n_0 every member of ν is of the form $F_{n_0} f$. We then say that ν is an n_0-cover. For any cover ν, we can write ν as $\nu = \bigvee_{i=1}^{I} \nu_i$ where each ν_i is a constant i-cover. We say that ν_i is the i-th *level* of ν, and that $\{\nu_i\}$ is the *decomposition of ν into levels.*

Finally, we say that a cover ν over E is *simple* if for each $f \in E$ there is exactly one n such that $F_n f \in \nu$. Any cover ν over E has a simple subcover over E.

The first step in proving that large covers ν exist such that (ν, ν^x) is ε-disjoint is the following lemma.

Lemma 4.11. *Let $0 < \varepsilon < \frac{1}{2}$ and $x \in \Omega$. Let ν be a cover over a set E. Then there exists a simple subcover $\mu \subseteq \nu$ over a set $E' = \{f_1, \ldots, f_I\} \subseteq E$ such that (μ, μ^x) is ε-disjoint and $|\cup \mu| \geq \frac{\varepsilon}{8+2C}|E|$.*

Proof. With regards the requirement that μ be simple, we merely note that we can replace ν with a simple subcover of ν and proceed from there.

We first prove the lemma under the assumption that ν is constant and then proceed to the general case.

Suppose ν is a constant n-cover over E. We will show that there exists $E' \subseteq E$ such that for $\mu = \nu|_{E'}$ we have (μ, μ^x) is ε-disjoint and $|\cup \mu| \geq \frac{\varepsilon}{8}|E|$. Let $E' = \{f_1, \ldots, f_I\} \subseteq E$ be a maximal set such that for $\mu = \nu|_{E'}$, (μ, μ^x) is ε-disjoint. Set $D = \cup(\mu \vee \mu^x)$. If $|D \cap E| > \frac{1}{2}|E|$ then by lemma 4.10,

$$|\cup \mu| \geq \frac{1-\varepsilon}{2}|D| \geq \frac{1}{4}|D| \geq \frac{1}{4}|D \cap E| > \frac{1}{8}|E|$$

so we are done. Otherwise $|D \cap E| \leq \frac{1}{2}|E|$. Then by minimality, for each $f \in E \setminus D$ it must hold that $|F_n f \cap D| \geq \varepsilon |F_n|$. Since each $g \in D$ can lie in at most $|F_n|$ right translates of F_n, we have $|D| \geq \frac{1}{|F_n|} \cdot \varepsilon |D \setminus E| \cdot |F_n| \geq \frac{\varepsilon}{2}|E|$, so

$$|\cup \mu| \geq \frac{1-\varepsilon}{2}|D| \geq \frac{(1-\varepsilon)\varepsilon}{4}|E| \geq \frac{\varepsilon}{8}|E|.$$

Returning to the case where the ν is not constant, let $\nu = \bigvee_{i=1}^{I} \nu_i$ be the levels of ν. Since ν is simple, the set of centers of ν_i and ν_j are disjoint fro $i \neq j$. We define inductively for $i = I$ down to $i = 1$ subcovers μ_i of ν_i over sets $E'_i = \{f_{i1}, \ldots, f_{iM(i)}\}$ such that (μ_i, μ_i^x) is ε-disjoint, as follows: Assume we have defined μ_j for $j > i$ and define

$$\nu'_i = \left\{F_i f \in \nu_i \,:\, F_i f \cap \left(\cup(\mu_j \vee \mu_j^x)\right) = \emptyset \text{ for every } j > i\right\}.$$

This is equivalent to saying that $\nu'_i = \nu_i|_{E_i}$, where

$$E_i = \operatorname{dom} \nu_i \setminus \bigcup_{j>i} \left(\cup^+(\mu_j \vee \mu_j^x)\right).$$

Now apply the one-level case proved above to ν'_i to obtain a subcover μ_i which is the restriction of ν_i to $E'_i = \{f_{i1}, \ldots, f_{iM(i)}\} \subseteq E_i$.

Finally, set $E' = \cup E'_i$, and order E' as in

$$E' = \{f_i\}_{i=1}^{M} = \{f_{I1}, \ldots, f_{IM(I)}, \ldots, f_{11}, \ldots, f_{1M(1)}\}$$

(note that the E'_i are pairwise disjoint), and define $\mu = \cup \mu_i = \nu|_{E'}$.

It is easy to verify from the construction that (μ, μ^x) is ε-disjoint. We must verify that $\cup\mu$ is large enough. Write $D = \cup^+(\mu \vee \mu^x)$ and note that $E_i \supseteq \operatorname{dom} \nu_i \setminus D$ so

$$\cup_i E_i \supseteq (\cup_i \operatorname{dom} \nu_i) \setminus D = E \setminus D.$$

Hence

$$|\cup \mu| = \sum_{i=1}^{I} |\cup \mu_i|$$

which by the definition of μ_i and the one-level case proved at the beginning implies

$$\begin{aligned} |\cup \mu| &\geq \frac{\varepsilon}{8}\sum_{i=1}^{I}|E_i| \\ &\geq \frac{\varepsilon}{8}|\bigcup_{i=1}^{I}E_i| \\ &\geq \frac{\varepsilon}{8}(|E|-|D|) \\ &\geq \frac{\varepsilon}{8}(|E|-\frac{2C}{1-\varepsilon}|\cup\mu|) \end{aligned}$$

the last inequality since by lemma 4.10 we have $|D| \leq \frac{2C}{1-\varepsilon}|\cup\mu|$. After rearranging this is

$$|\cup\mu| \geq (\frac{8}{\varepsilon}+\frac{2C}{1-\varepsilon})^{-1}|E| \geq \frac{\varepsilon}{8+2C}|E|. \qquad \square$$

Next, we want to get rid of the ε in the constant on the right hand side of the last inequality. For this, the idea is to apply the last lemma repeatedly to several “layers” of covers. If we arrange that the covers are invariant enough with respect to each other, we can at each stage obtain a subcover with size almost a fixed fraction of the size of what remains after the last stage, and repeating this enough times we will get a large fraction of E.

Definition 4.12. For fixed $\{F_n\}$, $\varepsilon > 0$, if $m < n$ are integers we write $m \prec_\varepsilon n$ if for all $n' \geq n$ $|(\cup_{i\leq m}F_i^{-1})F_{n'}| \leq (1+\varepsilon)|F_{n'}|$. For covers ν_1, ν_2 we write $\nu_1 \prec_\varepsilon \nu_2$ if $\max\nu_1 \prec_\varepsilon \min\nu_2$.

It is easy to verify that the relation $\prec_\varepsilon$ is transitive. As an example of the relation $\prec_\varepsilon$, note that if $\{F_n\}$ is $(1+\varepsilon)$-tempered, then $i \prec_\varepsilon (i+1)$ for all i. Also note that for any Følner sequence $\{F_n\}$ and for any m and ε, $m \prec_\varepsilon n$ is true for large enough n.

Definition 4.13. For $n \in \mathbb{N}$ and $\varepsilon > 0$, let

$$F_n^{(\varepsilon)} = (\bigcup_{i:i\prec_\varepsilon n} F_i^{-1})F_n.$$

For a cover ν let $\cup^\varepsilon\nu = \cup\{F_n^{(\varepsilon)}f \,:\, F_nf \in \nu\}$.

By definition of the $\prec_\varepsilon$ relation, $|F_n^{(\varepsilon)}| \leq (1+\varepsilon)|F_n|$. The following lemmas are proved like 4.9 and 4.10:

Lemma 4.14. *If ν is an ε-disjoint cover, then $|\cup^\varepsilon \nu| \leq \frac{1+\varepsilon}{1-\varepsilon}|\cup\nu|$, and if (ν,ν^x) is ε-disjoint then $|(\cup^\varepsilon\nu)\cup(\cup^\varepsilon\nu^x)| \leq \frac{2(1+\varepsilon)}{1-\varepsilon}|\cup\nu|$.*

Proposition 4.15. *Let $0 < \varepsilon < \frac{1}{2}$ and $x \in \Omega$. Then there exists an integer N such that for any sequence of N covers $\nu_1 \prec_\varepsilon \nu_2 \prec_\varepsilon \ldots \prec_\varepsilon \nu_N$ over E there a exists simple subcover $\mu \subseteq \bigvee_{k=1}^N \nu_k$ such that (μ, μ^x) is ε-disjoint and $|\cup\mu| \geq \frac{1}{8}|E|$.*

Proof. We will describe a subcover $\mu \subseteq \bigvee_{k=1}^N \nu_k$, and show that if N was large enough to begin with, the desired properties hold.

The proof is similar to lemma 4.11. Define inductively, from $i = N$ down to $i = 1$, subcovers μ_i of ν_i defined over sets $E_i' = \{f_{i1}, \ldots, f_{iM(i)}\} \subseteq E$ as follows: Assuming we have defined μ_j, E_j' for $j > i$, set

$$E_i = E \setminus \bigcup_{j \geq k} \left(\operatorname{dom}\mu_j \,\cup\, (\cup^\varepsilon(\mu_j \vee \mu_j^x))\right)$$

and $\nu_i' = \nu_i|_{E_i}$. μ_i and E_i' are obtained by and applying lemma 4.11 to ν_i'. Finally, define $E' = \cup_{i \leq I} E_i'$ and order E' as in lemma 4.11

$$E' = \{f_i\}_{i=1}^M = \{f_{I1}, \ldots, f_{IM(I)}, \ldots, f_{11}, \ldots, f_{1M(1)}\}$$

(as before the E_i' are pairwise disjoint). Set $\mu = \cup\mu_j$ It is easy to verify from the construction that μ is simple and (μ, μ^x) is ε-disjoint. We now show that $|\cup\mu| \geq \frac{1}{8}|E|$.

Set $D = \cup^\varepsilon(\mu \vee \mu^x)$. Then $|D| \leq \frac{2(1+\varepsilon)}{1-\varepsilon}|\cup\mu|$. Also, clearly $|\operatorname{dom}\mu| \leq |\cup\mu|$ because μ is ε-disjoint. Thus for each $i = 1, \ldots, N$,

$$\begin{aligned} |E_i| &\geq |E| - |D| - |\operatorname{dom}\mu| \\ &\geq |E| - (\frac{2(1+\varepsilon)}{1-\varepsilon} + 1)|\cup\mu \\ &\geq |E| - \frac{3-\varepsilon}{1-\varepsilon}|\cup\mu| \end{aligned}$$

so by lemma 4.11

$$\begin{aligned} |\cup\mu| &= \sum_{i=1}^N |\cup\mu_i| \geq \\ &\geq \sum_{i=1}^N \frac{\varepsilon}{8+2C}|E_i| \\ &\geq \sum_{i=1}^N \frac{\varepsilon}{8+2C}(|E| - \frac{3-\varepsilon}{1-\varepsilon}|\cup\mu|) \\ &\geq \frac{\varepsilon N}{8+2C}(|E| - \frac{3-\varepsilon}{1-\varepsilon}|\cup\mu|) \end{aligned}$$

and after rearranging this is

$$|\cup \mu| \geq \left(\frac{8+2C}{\varepsilon N} + \frac{(3-\varepsilon)}{(1-\varepsilon)} \right)^{-1} |E|$$

for N large enough this is greater than $\frac{1}{8}|E|$ $\square$

We will also need the following lemma, which can be proved by the same techniques as above. It was first proved by E. Lindenstrauss in [5], or see [11] for a different approach more similar to ours.

Lemma 4.16. *Let $0 < \varepsilon < \frac{1}{2}$. Then there exists an integer N such that for any sequence of N covers $\nu_1 \prec_\varepsilon \nu_2 \prec_\varepsilon \ldots \prec_\varepsilon \nu_N$ over E, there exists a simple subcover $\nu \subseteq \bigvee_{i=1}^N \nu_i$ such that ν is ε-disjoint and $|\cup \nu| \geq (1-2\varepsilon)|E|$.*

We end this section with a combinatorial result akin to lemma 4.2:

Lemma 4.17. *(Number of ε-disjoint covers) Let $L < N$ and $0 < \varepsilon < \frac{1}{2}$. Then the number of simple ε-disjoint covers ν with centers in F_N, $\min \nu > L$ and $\cup\nu \subseteq F_N$ is at most $2^{\rho(\varepsilon,L)|F_N|}$, where $\lim_{L\to\infty} \rho(\varepsilon, L) = 0$, uniformly in ε.*

Proof. Each cover $\nu = \left\{F_{n(i)} f_i\right\}_{i=1}^I$ of the type we are interested in is determined by its centers $\{f_1, \ldots, f_I\}$ and the values of $n(i)$ for $i = 1, \ldots, I$. If ν is ε-disjoint and $\cup\nu \subseteq F_N$, we have

$$|F_N| \geq |\cup \nu| \geq (1-\varepsilon) \sum_{i=1}^{I} |F_{n(i)} f_i| \geq (1-\varepsilon) \cdot I \cdot L$$

since $|F_n f| \geq n$ and $\min \nu > L$. Thus $I \leq \frac{1}{(1-\varepsilon)L}|F_N|$. Since we only consider simple covers, the number of possible sets of centers for such ε-disjoint covers ν with centers in F_N is bounded by the number of subsets of size at most $\frac{1}{(1-\varepsilon)a_L}|F_N|$ of $|F_N|$, which by lemma 4.2 is bounded above by

$$2^{\delta(\frac{1}{(1-\varepsilon)L})|F_N|}.$$

We now calculate how many ε-disjoint covers ν exist which satisfy our assumptions and have $\operatorname{dom} \nu = E$ for some fixed set E. We use a coding argument: Given $\nu = \left\{F_{n(i)} f_i\right\}_{i=1}^I$ as above, we write down the values $n(i)$ in binary notation in some pre-determined order (eg fix an order on G and write $n(i)$ down according to the order of the f_i). In order to decode the resulting concatenated sequence of 0s and 1s we must insert “punctuation” into the resulting concatenation of numbers. We therefore code each binary digit 0 as 00 and each 1 as 11, and terminate each number with the pair 01. Using this scheme, we can clearly recover ν from E and the codeword. In order to code $n(i)$ in this manner we use $2 \log n(i) + 2$ bits. Recalling that we always assume $|F_n| \geq n$, the length

of the resulting codeword is

$$\begin{aligned}\sum_{i=1}^{I}(2\log n(i)+2) &\leq \sum_{i=1}^{I}(2+2\log|F_{n(i)}f_i|)\\ &\leq 2\cdot I+2\sum_{i=1}^{I}|F_{n(i)}|\frac{\log|F_{n(i)}|}{|F_{n(i)}|}\\ &\leq \frac{2}{(1-\varepsilon)L}|F_N|+2\cdot\frac{\log L}{L}\sum_{i=1}^{I}|F_{n(i)}f_i|\\ &\leq \frac{2}{(1-\varepsilon)L}|F_N|+\frac{2\log L}{L}\cdot\frac{|\cup F_{n(i)}f_i|}{1-\varepsilon}\\ &\leq \frac{2}{(1-\varepsilon)L}|F_N|+\frac{2\log L}{(1-\varepsilon)L}|F_N|.\end{aligned}$$

Thus there are at most $2^{(\frac{2}{(1-\varepsilon)L)}+\frac{\log L}{(1-\varepsilon)L})|F_N|}$ such covers over a fixed set E.

Putting this all together, we see that the number of ε-disjoint covers ν with $\min\nu > L$ and $\cup\nu\subseteq F_N$ is at most

$$\exp\left(\ln 2\cdot\left(\frac{2\log L}{(1-\varepsilon)L}+\delta(\frac{1}{(1-\varepsilon)L})+\frac{2}{(1-\varepsilon)L}\right)|F_N|\right)$$

which proves the lemma. □

4.4. **Proof of the lower bound for T_* (Part II).**

Proof. (of theorem 4.3) Fix $\{F_n\}$,$\{W_n\}$ and the $(X_g)_{g\in G}$, defined on $(\Omega,\mathcal{F},P)$. Suppose $P(T_*^{(F,W)} < h-\varepsilon) > p > 0$. We are going to use lemma 4.15 to construct a sequence $\{c_n\}$ of F_n-codes such that eventually almost surely $x(F_n)$ can be reconstructed from $c_n(x)$, and

$$\frac{1}{|F_n|}\ell(c_n(x))\leq h-\frac{p}{5}\varepsilon.$$

This would contradict the fact that entropy is a lower bound for the achievable coding rates (section 4.1), since we can proceed as in lemma 4.1 and turn $\{c_n\}$ into a faithful code with the same rate.

Fix a $\eta > 0$ and a series $\{\Phi_n\}$ of sets of entropy-typical F_n-patterns, $\Phi_n\subseteq\Sigma^{F_n}$, such that for every $\varphi\in\Phi$,

$$2^{-(h+\eta)|F_n|}\leq P([\varphi])\leq 2^{-(h-\eta)|F_n|}$$

and $x\in[\Phi_n]$ eventually almost surely. For $x\in\Omega$ and N fixed, consider the pattern $x(F_N)$; the codeword $c_N(x)$ depends only on this pattern, and is constructed as follows:

Step 1: Consider the sets $F_k f \subseteq F_N$ such that $T_k(fx) < 2^{(h-\varepsilon/2)|F_k|}$, and from this collection try to extract a cover μ such that (μ, μ^x) is η-disjoint and very large. To be precise, we require that $|\cup \mu| \geq \frac{p}{8}|F_N|$. However, we try to choose μ so that the next step is possible:

Step 2: Consider the sets $F_k f \subseteq F_N$ for which $fx \in [\Phi_k]$, and from this collection try to extract an η-disjoint cover ν which is also disjoint from $\cup\mu$. We choose ν as large as possible; we require that $\cup\nu$ and $\cup\mu$ together cover all but a 3η-fraction of F_N.

Step 3: If we cannot find μ, ν as in steps 1,2 with the required sizes, we define $c_N(x)$ to be the empty word.

Step 4: Otherwise, let A_1 be the set guarenteed by lemma 4.7 when applied to μ, so $A_1 \subseteq \cup\mu$ and $|A_1| < 2\eta|\cup\mu| \leq 2\eta|F_N|$, and A_1 has the property that if pattern $x_1(A \cup (\cup\mu^x))$ is known and if the elements $f^{x,k}$ are known for every $F_k f \in \mu$ then $x(\cup\mu)$ can be deduced. Also, let $A_2 = F_N \setminus ((\cup\mu) \cup (\cup\nu))$.

Step 5: Now define (assuming steps 1 and 2 were successful) $c_n(x) = \gamma_1\gamma_2\gamma_3$, where

- γ_1 encodes the pattern $x(A_1 \cup A_2)$.
- γ_2 encodes the pattern $x(\cup\nu)$.
- γ_3 encodes the cover μ along with the elements $f^{x,k}$ for each $F_k f \in \mu$.

It is clear from the definition of A_1 that if steps 1,2 succeeded then we can reconstruct $x(F_N)$ from $c_N(x)$. It remains to show that for almost every x, steps 1 and 2 will succeed for large enough N, and that in this case the encoding, as described in step 5, can be carried out in such a way that $\ell(c_N(x)) \leq (h - p/10)|F_N|$.

We begin by showing that for large enough N, steps 1 and 2 will succeed. Fix integers M, L to be determined later, and select a sequence of intervals $\{[a_i; b_i]\}_{i=1}^{2L}$ such that $a_1 = M$, $b_i \prec_\eta a_{i+1}$ and

(1) For a set of $x \in \Omega$ with probability greater than p there exist indices $m_i(x) \in [a_i; b_i]$ such that $T_{m_i(x)}(x) \leq h - \varepsilon$ for each i.

(2) For a set of $x \in \Omega$ with probability greater than $1 - \eta$ there exist indices $n_i(x) \in [a_i; b_i]$ such that $x \in \Phi_{n_i(x)}$ for each i.

Such a sequence is easy to construct using the ergodic theorem, the fact that $P(T_* \leq h - \varepsilon) > p$, and our choice of $\{\Phi_n\}$; we omit the details.

From the pointwise ergodic theorem and the asymptotic invariance of $\{F_n\}$, for almost every $x \in \Omega$ it holds for large enough N that

$$\frac{1}{|F_N|}\#\left\{ f \in F_N \;\middle|\; \begin{array}{c} \text{for } L+1 \leq i \leq 2L \text{ there exist indices} \\ m_i(f) \in [a_i; b_i] \text{ with } T_{m_i(f)}(fx) \leq h - \varepsilon \\ F_{m_i(f)} f \subseteq F_N \text{ and } (F_{m_i(f)} f)^x \subseteq F_N \end{array} \right\} > p \tag{4.1}$$

$$\frac{1}{|F_N|}\#\left\{ f\in F_N \;\middle|\; \begin{array}{c}\text{for } 1\leq i\leq L \text{ there exist indices}\\ n_i(f)\in[a_i;b_i] \text{ with}\\ fx\in\Phi_{n_i(f)} \text{ and } F_{n_i(f)}f\subseteq F_N\end{array}\right\} > 1-\eta. \tag{4.2}$$

For x, N such that the above holds, let $m_i(f)$ and $n_i(f)$ be as in (4.1) and (4.2).

Let $E_1 \subseteq F_N$ be the set described in 4.1, and $\mu_i = \left\{F_{m_i(f)}f\right\}_{f\in E_1}$, so we have covers μ_i over E_1. Note that $\mu_1 \prec_\eta \mu_2 \prec_\eta \ldots \prec_\eta \mu_L$ by our choice of $[a_i;b_i]$; so assuming L was chosen large enough we can apply lemma 4.16 to $\mu_1,\ldots,\mu_L$ and obtain a simple subcover $\mu \subseteq \bigvee_{i=1}^L \mu_i$ over some set $E_1' = \{f_1,\ldots,f_I\} \subseteq E_1$ such that (μ,μ^x) is η-disjoint and $|\cup\mu| \geq \frac{1}{8}|E_1|$, which means that

$$|\cup\mu| \geq \frac{p}{8}|F_N|.$$

Let $E_2 \subseteq F_N$ be the set described in 4.2, and $\nu_i = \left\{F_{n_i(f)}f\right\}_{f\in E_2}$, so we have covers ν_i over E_2. As with the μ_i, we have $\nu_1 \prec_\eta \nu_2 \prec_\eta \ldots \prec_\eta \nu_L$, and furthermore $\nu_L \prec_\eta \mu$. Let ν_i' be the restriction of ν_i to $E_2 \setminus \cup^\eta \mu$, so each member set of ν_i' is disjoint from $\cup\mu$. Assuming L is large enough we apply lemma 4.16 and obtain an η-disjoint simple subcover $\nu \subseteq \bigvee_{i=1}^L \nu_i'$ over some set $E_2' \subseteq E_2 \setminus (\cup^\eta \mu)$ such that $|\cup\nu| \geq (1-2\eta)|E_2 \setminus (\cup^\eta\mu)|$. We then have

$$|\cup\nu| \geq (1-2\eta)\left(|E_2| - |\cup^\eta \mu|\right) \geq (1-2\eta)\left((1-\eta)|F_N| - \frac{1+\eta}{1-\eta}|\cup\mu|\right).$$

By construction μ,ν are disjoint, and putting the last two paragraphs together, we have

$$|\cup(\mu\vee\nu)| = |\cup\mu| + |\cup\nu| \geq$$
$$\geq (1-(1-2\eta)\frac{1+\eta}{1-\eta})|\cup\mu| + (1-2\eta)(1-\eta)|F_N| \geq (1-3\eta)|F_N|.$$

Thus all but a 3η-fraction of F_N is covered by $\mu\cup\nu$, and we have shown that for almost every x, for large enough N there exist μ,ν as required by steps 1 and 2 of the coding construction.

We complete the proof by giving a more detailed description of the encoding of $c_N(x)$ (in the case where steps 1,2 succeeded) by specifying exactly how to construct $\gamma_1,\gamma_2,\gamma_3$, and show that the codeword length is indeed bounded by $(h-p/5)|F_N|$.

Recall the functions $\delta(\cdot)$ and $\rho(\cdot,\cdot)$ defined in lemmas 4.2 and 4.17 respectively. In the estimates below we ignore rounding errors; the reader may verify that these may be taken into account without disrupting the proof.

Fix x, N so that steps 1,2 of the construction succeeded. Let μ,ν,A_1,A_2 be as in the construction.

Write $H = A_1\cup A_2$. The word γ_1 will first encode the set H and then the values $x(f)$ for $f\in H$. The former takes $\delta(5\eta)|F_N|$ bits, because by lemma 4.2 we can costruct a list of all $2^{\delta(5\eta)|F_n|}$ subsets of F_n with size at most $5\eta|F_n|$ and then describe H by giving its index in the list. Next, using some

fixed ordering of G (which induces an order on H), we write down the symbols $x(f)$ for $f \in H$. For this we need an additional $|H| \log \Sigma \leq (5\eta \log \Sigma)|F_N|$ bits. So all together $x(H)$ can be coded in $(\delta(5\eta) + 5\eta \log \Sigma)|F_N|$ bits.

The word γ_2 will be coded by specifying first the cover ν and then the values $x(F_k f)$ for $F_k f \in \nu$. The first task is accomplished using $\rho(\eta, M)|F_N|$ bits by lemma 4.17. Now for each $F_k f \in \nu$, according to some fixed order, we record the pattern $x(F_k f)$ by giving the index of the pattern $x(F_k f)$ in Φ_k. This requires $(h+\eta)|F_k|$ bits by our choice of $\{\Phi_n\}$, because $|\Phi_k| \leq 2^{(h+\eta)|F_k|}$, and because by the definition of ν, $fx \in [\Phi_k]$ whenever $F_k f \in \nu$. Thus to describe all the patterns $x(F_k f)$ for $F_k f \in \nu$, we need

$$\sum_{\nu} (h+\eta)|F_k f| = (h+\eta) \sum_{\nu} |F_k f| \leq \frac{h+\eta}{1-\eta} |\cup \nu|$$

bits. All together, we see that we can code γ_2 with at most $\rho(\eta, M)|F_N| + \frac{h+\eta}{1-\eta}|\cup \nu|$ bits.

The word γ_3 is coded by first describing μ and then, for each $F_k f \in \mu$, we describe $f^{x,k}$. The former is again achieved with $\rho(\eta, M)|F_N|$ bits (the same calculation as above applies), while the latter requires at most $(h-\varepsilon)|F_{\mu(f)}|$ bits for each $F_k f \in \mu$, by definition of $f^{x,k}$. Note that this is where we use the assumption that $\{W_n\}$ is increasing: for if it were not, then in order to record $f^{x,k}$ we would need to specify also which set W_n it is from. In other words, we would need to record $R_k(fx)$, and this could be expensive (there is no upper bound on $R_k(fx)$). However, since $\{W_n\}$ is increasing, we can assume that $f^{x,k}$ comes from the largest set W_n satisfying $|W_n| < 2^{(h-\varepsilon)|F_k|}$). Thus estimating as we did for γ_2 we find that γ_3 can be coded in

$$\rho(\eta, M)|F_N| + (h-\varepsilon) \sum_{\mu} |F_k f| \leq \rho(\eta, M)|F_N| + \frac{h-\varepsilon}{1-\eta} |\cup \mu|$$

bits.

Putting all this together, we find

$$\begin{aligned}
\ell(c_N(x)) &= \ell(\gamma_1) + \ell(\gamma_2) + \ell(\gamma_3) \\
&\leq \left\{ \begin{array}{c} (\delta(7\eta) + 5\eta \log \Sigma)\,|F_N| + \\ + \left(\rho(\eta, M)|F_N| + \frac{h+\eta}{1-\eta}|\cup \nu|\right) \\ + \left(\rho(\eta, M)|F_N| + \frac{h-\varepsilon}{1-\eta}|\cup \mu|\right) \end{array} \right\} \\
&\leq \left\{ \begin{array}{c} (\delta(7\eta) + 5\eta \log \Sigma + 2\rho(\eta, M))\,|F_N| + \\ + \frac{h+\eta}{1-\eta}\,|(\cup \nu) \cup (\cup \mu)| - \frac{\varepsilon-\eta}{1-\eta}|\cup \mu| \end{array} \right\} \\
&\leq (h + \tau(\eta, M))|F_N| - \frac{\varepsilon - \eta}{1-\eta} \cdot \frac{p}{8}|F_N|
\end{aligned}$$

where $\tau(\eta, M) \to 0$ as $M \to \infty$ uniformly in η. So for M large enough and η small enough we have

$$\ell(c_N(x)) \leq (h - \frac{p}{10}\varepsilon)|F_N|.$$

This completes the proof. $\square$

4.5. **The lower bound for U_*.** In this section we prove that

$$U_*^{(F,W)}(x) = \liminf_{k\to\infty} \inf_n -\frac{1}{|F_k|}\log(U_{k,n}^{(F,W)}(x) - \frac{1}{|W_n|}) \geq h$$

almost surely if $\{F_n\}$,$\{W_n\}$ satisfy the conditions of theorem 1.11.

As was already mentioned in the introduction, the correction term $-\frac{1}{|W_n|}$ in the definition of U_* is necessary because $U_{k,n}$ counts the central pattern $x(F_k)$; without it, if for instance $1_G \in W_1$, we would have for every k that $U_{k,1} \geq \frac{1}{|W_1|}$, and this implies $U_* = 0$ for any process.

In the previous section, we defined the quantities $V_{k,n}$ and V^*, and showed that $U^* = V^*$. We may analogously define

$$V_* = \liminf_k \inf_n -\frac{1}{|F_k|}\log(V_{k,n} - \frac{1}{|W_n|})$$

as a version of U_* which counts all repetitions of $x(F_k)$ in $x(W_n)$, but we cannot, as was the case with V^*, show that $U_* = V_*$, because when the correction factor is taken into account, if we compare $U_{k,n} - \frac{1}{|W_n|}$ with $V_{k,n} - \frac{1}{|W_n|}$ as in equation 3.1, we find that all we can say is

$$U_{k,n} - \frac{1}{|W_n|} \leq V_{k,n} - \frac{1}{|W_n|} \leq |F_k|^2(U_{k,n} - \frac{1}{|W_n|}) + \frac{|F_k|^2}{|W_n|}$$

and this gives us $U_* \geq V_*$ but not the reverse inequality.

Furthermore, the fact that $U_{k,n}$ does not count repetitions of $x(F_k)$ which intersect F_k is necessary for our proof that $U_* \geq h$; the proof cannot be adapted to the case of V_*. The importance of the requirement that $U_{k,n}$ not count the repetition which intersect the original pattern stems from the following observation, which is used in an essential way in the proof: For almost every x, for large enough k, $U_{k,n}(x) = 0$ unless $|W_n| > 2^{(h-\varepsilon)|F_k|}$. This follows from the fact that if $U_{k,n}(x) > 1$ then there is within W_n a repetition of $x(F_k)$ *disjoint* from F_k, and so $R_k(x) \leq n$, which in turn implies $T_k(x) \leq |W_n|$ (we will assume that $\{W_i\}$ is increasing); now use the fact that for large enough k, by theorem 4.3, $T_k(x) > 2^{(h-\varepsilon)|F_k|}$. In the case of $V_{k,n}$, if $V_{k,n} > 1$ it may be because of a repetition of $x(F_k)$ which intersects F_k, and this argument fails. Thus we can't allow repetitions which intersect F_k to be counted in $U_{k,n}$ because R_k was defined as the first repetition of $x(F_k)$ *disjoint* from F_k. If, however, we were able to prove an analogue of $T_* \geq h$ (theorem 4.3) for the case where R_k counts *all* repetitions in W_n of $x(F_k)$, including those intersecting F_k, then the proof below would work just as

well for V_*. This version of theorem 4.3 is true, for instance, in the case of $W_n = [-k_n; k_n]^d \subseteq \mathbb{Z}^d$. This was proved in [8].

One can, however, define

$$\hat{V}_{k,n}(x) = \frac{1}{|W_n|} \max \left\{ |E| \;\middle|\; \begin{array}{c} 1_G \in E \subseteq W_n \text{ and if } 1_G \neq f \in E \text{ then} \\ F_k \cap F_k f = \emptyset \,, \text{ and } fx \in [x(F_k)] \end{array} \right\}.$$

Then we have that

$$U_{k,n} - \frac{1}{|W_n|} \leq \hat{V}_{k,n} - \frac{1}{|W_n|} \leq |F_k|^2 (U_{k,n} - \frac{1}{|W_n|})$$

and so $\hat{V}_* = \liminf_{k\to\infty} \inf_n -\frac{1}{|F_k|} \log(\hat{V}_{k,n}(x) - \frac{1}{|W_n|}) = U_*$.

Our proof that $U_* \geq h$ is based on the bound $T_* \geq h$. It is not hard to see that $U_*^{(F,W)} \geq h$ implies $T_*^{(F,W)} \geq h$. We would like to show the reverse, namely, that $U_*^{(F,W)} < h$ implies $T_*^{(F,W)} < h$. The idea behind the proof is that from the relation $U_{k,n}(x) - \frac{1}{|W_n|} > 2^{-(h-\varepsilon)|F_k|}$, which means that there are "too many" repetitions of $x(F_k)$ in W_n, we can sometimes deduce that many of the points $g \in W_n$ at which $x(F_k)$ repeats are such that another repetition of the same pattern appears "too close" to g. If we could show this, we can appeal to the fact that $T_* \geq h$ to obtain a contradiction. One problem here is that the window sets W_n may be unsuitable for capturing the idea of "too close". This will happen if the sequence $\{W_n\}$ grows too quickly, and in this case we will not be able to obtain a contradiction through $T_k^{(F,W)}$. We therefore introduce a second window sequence $\{Y_n\}$, which will grow slowly enough that with respect to $\{Y_n\}$, a drop of $T_k^{(F,Y)}$ below entropy can be observed. We will show that for a nice enough sequence $\{Y_n\}$, if with positive probability $U_*^{(F,W)} < h - \varepsilon$ is true, then it is not almost surely true that $T_*^{(F,Y)} \geq h$, which is impossible.

We recall the following definitions, which were given in the introduction:

Definition. Let $\{F_n\}, \{W_n\}$ be sequences of finite subsets of G. An increasing sequence $\{Y_n\}$ of finite subsets of G is called an *interpolation sequence for* $\{F_n\}, \{W_n\}$ if

(1) $\{Y_n\}$ is filling and increasing .
(2) If $|W_n| \geq |Y_m|$ then $|Y_m W_n| \leq C|W_n|$ for some constant C.
(3) For every pair of real numbers $0 \leq \alpha < \beta$, for every large enough n there is an index k such that $2^{\alpha|F_n|} \leq |Y_k| \leq 2^{\beta|F_n|}$.

The main lemma we will need about interpolation sequences is that too frequent recurrence of $x(F_k)$ in $x(W_n)$ implies that $T_k^{(F,Y)}(fx)$ is too small for many $f \in W_n$.

Lemma 4.18. *Let Y be $1/C$-filling and $E \subseteq W \subseteq G$ finite sets. Suppose that $|YW| \leq C|W|$. If $\frac{|E|}{|W|} \geq \alpha$ then*

$$\frac{1}{|W|} |\, \{f \in E \,:\, |Yf \cap E| > \beta\} \,| \geq \alpha - C^2 \frac{\beta}{|Y|}.$$

Proof. Let $E_- = \{f \in E : |Yf \cap E| \leq \beta\}$, and let $\{Yf_i\}_{i=1}^I$ be an incremental cover of E_- with $f_i \in E_-$. Then by the basic lemma 3.1 we have that

$$\frac{|E_-|}{|\cup_i Yf_i|} \leq C\frac{\beta}{|Y|}$$

since $\cup_i Yf_i \subseteq YW$ we have

$$|E_-| \leq C\frac{\beta}{|Y|}|YW| \leq C^2\frac{\beta}{|Y|}|W|$$

so

$$\frac{1}{|W|}|\left\{f \in E \,:\, \frac{|Yf \cap E|}{|Y|} > \beta\right\}| = \frac{1}{|W|}(|E| - |E_-|) \geq \alpha - C^2\frac{\beta}{|Y|}. \qquad \square$$

Corollary 4.19. *Let $0 < \delta < \varepsilon/2$, and let $\{Y_n\}$ be an interpolation sequence for $\{F_n\}, \{W_n\}$. If for some k, n and $x \in \Omega$ we have $U_{k,n}^{(F,W)}(x) \geq 2^{(h-\varepsilon)|F_k|}$, and if $|W_n| > 2^{(h-\delta)|F_k|}$, then if k is large enough,*

$$\frac{1}{|W_n|}|\left\{f \in W_n \,\middle|\, \begin{array}{c} fx \in [x(F_k)] \text{ and} \\ T_k^{(F,Y)}(fx) < 2^{(h-\varepsilon/2)|F_k|} \end{array}\right\}| > \frac{1}{2}2^{-(h-\varepsilon)|F_k|}.$$

Proof. Set $E = \{f \in W_n \,:\, fx \in [x(F_k)]\}$ and $\beta = |F_k|^2$. For k large enough there exists an m such that

$$2^{(h-2\varepsilon/3)|F_k|} \leq |Y_m| \leq 2^{(h-\varepsilon/2)|F_k|}$$

Since $|W_n| > 2^{(h-\delta)|F_k|}$, we have $|Y_mW_n| \leq C|W_n|$, so by the lemma (with $\alpha = 2^{-(h-\varepsilon)|F_k|}$)

$$\frac{1}{|W_n|}|\left\{f \in E \,:\, |Y_mf \cap E| > |F_k^2|\right\}| \geq$$

$$\geq 2^{-(h-\varepsilon)|F_k|} - C^2|F_k|^2 2^{-(h-2\varepsilon/3)|F_k|} \geq \frac{1}{2}2^{-(h-\varepsilon)|F_k|}$$

for k large enough. Now note that if for some $f \in E$ we have $|Y_mf \cap E| > |F_k|^2$ then at least one of the repetitions of $fx(F_k)$ within $fx(Y_m)$ is disjoint from $fx(F_k)$, so $T_k^{(F,Y)}(fx) \leq |Y_m| < 2^{(h-\delta)|F_k|}$. This completes the proof. $\square$

The second requirement we make of $\{W_n\}$ is that it be incompressible. We recall the definition:

Definition. An increasing sequence $\{W_n\}$ of finite subsets of G with $1_G \in W_n$ is said to be *incompressible with constant C* (or C-incompressible) if for any incremental sequence $\{W_{n(i)}f_i\}$, the number of the sets $W_{n(i)}f_i$ containing 1_G is at most C.

An incompressible sequence $\{W_n\}$ is filling; if $\{W_n\}$ is incompressible with constant C then it will be filling with constant $\frac{1}{C}$, since for a collection $\{W_{n(i)}f_i\}_{i\in I}$ as above, each $g \in \cup_{i\in I}F_{n(i)}f_i$ belongs to at most C of the sets $W_{n(i)}f_i$. Incompressible sequences are superior to filling sequences because of the following observation: If $\{W_n\}$ is incompressible and $\{W_{n(i)}f_i\}_{i=1}^I$ is incremental,

then for any finite $A \subseteq G$, we have

$$|A \cap \bigcup_{i \in I} W_{n(i)} f_i| \geq \frac{1}{C} \sum_{1 \leq i \leq I} |A \cap W_{n(i)} f_i|.$$

This is because every $a \in A$ is counted at most C times in the sum on the right.

The property of incompressibility is equivalent to the following property, which may be described as being filling relative to arbitrary subsets of G:

Lemma 4.20. *Let $\{W_n\}$ be incompressible, $H_1, H_2 \subseteq G$ and suppose $\left\{W_{n(i)} f_i\right\}_{i=1}^{I}$ is an incremental sequence such that for each $i = 1, \ldots, I$ we have that*

$$\alpha \leq \frac{|H_1 \cap W_{n(i)} f_i|}{|H_2 \cap W_{n(i)} f_i|} \leq \beta.$$

Then

$$\frac{\alpha}{C} \leq \frac{|H_1 \cap (\cup_i W_{n(i)} f_i)|}{|H_2 \cap (\cup_i W_{n(i)} f_i)|} < C\beta$$

(this remains true for $\alpha = 0$ or $\beta = \infty$) .

It is not difficult to see that if $\{W_n\}$ satisfies the conclusion of the lemma then it is incompressible.

Proof. We show for instance the lower bound.

$$\frac{|H_1 \cap (\cup_i W_{n(i)} f_i)|}{|H_2 \cap (\cup_i W_{n(i)} f_i)|} \geq \frac{\frac{1}{C} \sum |H_1 \cap W_{n(i)} f_i|}{\sum |H_2 \cap W_{n(i)} f_i|} \geq$$

(we use here the fact that $\{W_n\}$ is $\frac{1}{C}$-filling)

$$\geq \frac{1}{C} \sum_i \frac{|H_2 \cap W_{n(i)} f_i|}{\sum_j |H_2 \cap W_{n(i)} f_j|} \cdot \frac{|H_1 \cap W_{n(i)} f_i|}{|H_2 \cap W_{n(i)} f_i|} \geq \frac{\alpha}{C}.$$

□

Theorem 4.21. *Let F_n be a tempered Følner sequence, $\{X_g\}_{g \in G}$ an ergodic process with entropy h. Let $\{W_n\}$ be an incompressible sequence. If there exists an interpolation sequence for $\{F_n\}, \{W_n\}$. Then*

$$\liminf_{k \to \infty} \inf_n -\frac{1}{|F_k|} \log U_{k,n}^{(F,W)}(x) \geq h.$$

Proof. Suppose the theorem is false. Then there is a set $B \subseteq \Omega$ with $P(B) > p > 0$ such that for every $x \in B$,

$$\limsup_{k \to \infty} \sup_n -\frac{1}{|F_k|} \log U_{k,n}^{(F,W)}(x) \leq h - \varepsilon$$

for some $\varepsilon > 0$. Let $\{\Phi_n\}$ be a sequence of sets of words, $\Phi_n \subseteq \Sigma^{F_n}$, such that

$$\frac{1}{|F_n|} \log |\Phi_n| \to h$$

and $x \in [\Phi_n]$ eventually almost surely. Fix L very large; we may assume that $x \in [\Phi_n]$ for every for $n \geq L$ and $x \in B$. We may also assume that L was chosen large enough that $|\Phi_k| \leq 2^{(h+\delta)|F_k|}$ for $k > L$.

For every $x \in B$ there exists an index $k(x) > L$ and an index $n(x)$ such that

$$U_{k(x),n(x)}(x) \geq 2^{-(h-\varepsilon)|F_{k(x)}|}$$

and without loss of generality we may assume that for every $x \in B$ we have $n(x) \leq M$ for some integer M. We further assume, increasing L if necessary, that for every $x \in B$, $|W_{n(x)}| > 2^{(h-\varepsilon/4)|F_{k(x)}|}$, because $T_*^{(F,W)} \geq h$ almost surely.

Let $x \in \Omega$ be typical in the sense that the ergodic theorem holds for B and every $[\Phi_k]$, $L \leq k \leq M$. By the ergodic theorem, we can choose N large enough so that

$$\frac{1}{|F_N|} |\{f \in F_N \,:\, fx \in B\}| > p.$$

For $f \in F_N$ such that $fx \in B$, we will write for brevity $k(f) = k(fx)$ and $n(f) = n(fx)$.

Step 1: Fix $L \leq k \leq M$ and $\varphi \in \Phi_k$, and let

$$H = \{f \in F_N \,:\, fx \in [\varphi]\} \ , \ E = \{f \in H \,:\, fx \in B \,, k(f) = k\} \,.$$

If $f \in E$ then $U_{k,n(f)}^{(F,W)}(fx) \geq 2^{-(h-\varepsilon)|F_k|}$. Some of the points $g \in W_{n(f)}f$ at which φ repeats will belong to E as well, and some won't. Write E^+ for the set of $f \in E$ for which "many" of the repetitions of φ in $W_{n(f)}f$ are in E:

$$E^+ = \left\{ f \in E \,:\, \frac{1}{|W_n|} |W_{n(f)} f \cap E| \geq 2^{-(h-\varepsilon/2)|F_k|} \right\}$$

and

$$E^- = E \setminus E^+ = \left\{ f \in E \,:\, \frac{1}{|W_n|} |W_{n(f)} f \cap E| < 2^{-(h-\varepsilon/2)|F_k|} \right\}.$$

We are interested in the elements of F_N for which $T_k^{(F,Y)} < 2^{-(h+\varepsilon/4)|F_k|}$. Let

$$D = \left\{ f \in H \,:\, T_k^{(F,Y)}(fx) < 2^{-(h+\varepsilon/2)|F_k|} \right\}.$$

Now if $f \in E^+$, then by lemma 4.19 we have that

$$|(E \cap W_{n(f)} f) \cap D| \geq \frac{1}{2} |E \cap W_{n(f)} f|.$$

Select an incremental cover $\left\{W_{n(f_i)}f_i\right\}_{i=1}^{I}$ of E^+ with $f_i \in E^+$. For each i we have that the relative density of $D \cap E$ in $W_{n(f_i)}f_i$ is at least half the density of E, so from lemma 4.20 we see that

$$\frac{|(E \cap D) \cap (\cup W_{n(f_i)}f_i)|}{|E \cap (\cup W_{n(f_i)}f_i)|} \geq \frac{1}{2C} \tag{4.3}$$

and therefore, since $E^+ \subseteq \cup W_{n(f_i)}f_i$, we have

$$|E \cap D| \geq \frac{1}{2C}|E^+|. \tag{4.4}$$

On the other hand, if $f \in E^-$, we have that the number of elements of H in $W_{n(f)}f$ is at least $2^{(\varepsilon/2)|F_k|}$ times the number of elements of E in $W_{n(f)}f$, so

$$|W_{n(f)}f \cap D| \geq \frac{1}{2}|W_{n(f)}f \cap H| \geq \frac{1}{2}2^{(\varepsilon/2)|F_k|} \cdot |E \cap W_{n(f)}f|.$$

Now select an incremental cover $\left\{W_{n(f_i)}f_i\right\}_{i=1}^{I}$ of E^-. From lemma we have

$$\frac{|D \cap (\cup W_{n(f_i)}f_i)|}{|E \cap (\cup W_{n(f_i)}f_i)|} \geq \frac{1}{2C}2^{(\varepsilon/2)|F_k|} \tag{4.5}$$

and therefore, since $E^- \subseteq \cup W_{n(f_i)}f_i$, we have

$$|D| \geq \frac{1}{2C}2^{(\varepsilon/2)|F_k|}|E^-|. \tag{4.6}$$

We now face two alternatives:

(1) If $|E^+| \geq \frac{1}{2}|E|$ then equation 4.4 gives us

$$|E \cap D| \geq \frac{1}{2C}|E^+| \geq \frac{1}{4C}|E|.$$

(2) Otherwise, if $|E_k^{\varphi,-}| > \frac{1}{2}|E_k^{\varphi}|$, equation 4.6 gives

$$|D| \geq \frac{1}{2C}2^{(\varepsilon/2)|F_k|}|E^-| \geq \frac{1}{4C}2^{(\varepsilon/2)|F_k|}|E|.$$

Step 2: For fixed k and $\varphi \in \Sigma^{F_k}$ write D^φ, E^φ for the sets D, E of step 1, thus making explicit the dependence on φ which was previously suppressed. Note that the collections $\left\{D^\varphi\right\}_{\varphi\in\Phi_k}$, $\left\{E^\varphi\right\}_{\varphi\in\Phi_k}$ are pairwise disjoint. For each $\varphi \in \Phi_k$ we have by the above that either (a) $|E^\varphi \cap D^\varphi| \geq \frac{1}{4C}|E^\varphi|$ or (b) $|D^\varphi| \geq \frac{1}{4C}2^{(\varepsilon/2)|F_k|}|E^\varphi|$. Let $\Phi_k^+ \subseteq \Phi_k$ be the set of φ's for which the first alternative holds, and Φ_k^- its complement in Φ_k. There are again two alternatives:

(1) $|\cup_{\varphi\in\Phi_k^+} E^\varphi| \geq \frac{1}{2}|\cup_{\varphi\in\Phi_k} E^\varphi|$. In this case

$$|\bigcup_{\varphi\in\Phi_k}(E^\varphi \cap D^\varphi)| \geq |\bigcup_{\varphi\in\Phi_k^+}(E^\varphi \cap D^\varphi)| =$$

$$= \sum_{\varphi\in\Phi_k^+} |E^\varphi \cap D^\varphi| \geq \frac{1}{4C} \sum_{\varphi\in\Phi_k^+} |E^\varphi| =$$

$$= \frac{1}{4C} |\bigcup_{\varphi\in\Phi_k^+} E^\varphi| \geq \frac{1}{8C} |\bigcup_{\varphi\in\Phi_k} E^\varphi|.$$

(2) $|\cup_{\varphi\in\Phi_k^-} E^\varphi| > \frac{1}{2}|\cup_{\varphi\in\Phi_k} E^\varphi|$. In this case

$$|\bigcup_{\varphi\in\Phi_k} D^\varphi| \geq |\bigcup_{\varphi\in\Phi_k^-} D^\varphi| =$$

$$= \sum_{\varphi\in\Phi_k^-} |D^\varphi| \geq \frac{1}{4C} 2^{(\varepsilon/2)|F_k|} \sum_{\varphi\in\Phi_k^-} |E^\varphi| =$$

$$\geq \frac{1}{4C} 2^{(\varepsilon/2)|F_k|} |\bigcup_{\varphi\in\Phi_k^+} E^\varphi| \geq \frac{1}{8C} 2^{(\varepsilon/2)|F_k|} |\bigcup_{\varphi\in\Phi_k} E^\varphi|.$$

Step 3: We now let k vary between L and M. Write $E_k = \cup_{\varphi\in\Phi_k} E^\varphi$ and $D_k = \cup_{\varphi\in\Phi_k} D^\varphi$. Note that

$$E_k = \{f \in F_N \,:\, fx \in B \text{ and } k(f) = k\}$$
$$D_k = \left\{f \in F_N \,:\, fx \in [\Phi_k] \text{ and } T_k^{(F,Y)}(fx) < 2^{(h-\varepsilon/2)|F_k|}\right\}$$

so $\{E_k\}_{L\leq k\leq M}$ are pairwise disjoint, but $\{D_k\}_{L\leq k\leq M}$ need not be. We saw in step (2) that for each $L \leq k \leq M$, either (a) $|E_k \cap D_k| \geq \frac{1}{8C}|E_k|$ or (b) $|D_k| \geq \frac{1}{8C} 2^{(\varepsilon/2)|F_k|}|E_k|$. Write

$$J = \left\{k \,:\, L \leq k \leq M \text{ and } |D_k| \geq \frac{1}{8C} 2^{(\varepsilon/2)|F_k|} |E_k|\right\}.$$

Let $k \in J$. If $|E_k| \geq 16C \cdot 2^{-(\varepsilon/2)|F_k|}|F_N|$ we would have

$$|D_k| \geq \frac{1}{8C} 2^{(\varepsilon/2)|F_k|} |E_k| \geq 2|F_N|.$$

However, $D_k \subseteq (\cup_{L\leq k\leq M} W_k)\,|F_N|$ and the latter set, if N is large enough, has size strictly less than $2|F_N|$, so the inequality $|D_k| \geq 2|F_N|$ is impossible. Thus for N large enough, if $k \in J$ it must be that $|E_k| \leq 16C \cdot 2^{-(\varepsilon/2)|F_k|}|F_N|$, so

$$\frac{1}{|F_N|} |\bigcup_{k\in J} E_k| = \frac{1}{|F_N|} \sum_{k\in J} |E_k| \leq \frac{16C}{|F_n|} \sum_{k=L}^{M} 2^{-(\varepsilon/2)|F_k|} \leq \frac{p}{2C|F_N|}$$

for L large enough. Therefore,

$$\frac{1}{|F_N|} |\bigcup_{L\leq k\leq M} D_k| \geq \frac{1}{|F_N|} |\bigcup_{k\notin J} E_k \cap D_k| = \frac{1}{|F_N|} \sum_{k\notin J} |E_k \cap D_k| \geq$$

$$\geq \frac{1}{8C|F_N|}\sum_{k\notin J}|E_k| = \frac{1}{8C|F_N|}|\bigcup_{k\notin J}E_k| =$$
$$= \frac{1}{8C|F_N|}|\bigcup_{L\leq k\leq M}E_k| - \frac{1}{8C|F_N|}|\bigcup_{k\in J}E_k| \geq$$

by assumption, $\frac{1}{|F_N|}\left|\bigcup_{L\leq k\leq M}E_k\right| \geq p$, so

$$\geq \frac{p}{8C} - \frac{p}{16C} = \frac{p}{16C}.$$

But this is impossible, since it implies that at least a $\frac{p}{16C}$-fraction of $f \in F_N$ are such that $\inf_{L\leq k\leq M} T_k^{(F,Y)}(fx) < 2^{(h-\varepsilon/2)|F_k|}$. Since L was arbitrarily large, this contradicts the fact that $T_*^{(F,Y)} \geq h$ a.s.. This completes the proof. □

In order to prove the theorem for $\{W_n\}$ a quasi-incompressible sequence (definition 1.13) we need an analogue of lemma 4.20. This lemma is used in the proof in step (1) to justify equations 4.3 and 4.5. Recalling what occurred there, we had fixed k and $\varphi \in \Sigma^{F_k}$ and considered a set $E \subseteq F_N$ of points f such that $U_{k,n(f)}^{(F,W)}(fx) < 2^{-(h-\varepsilon)|F_k|}$. Note that the right-hand side of this inequality depends only on k. Furthermore, one of our assumptions was that $|W_{n(f)}| > 2^{(h-\varepsilon/2)|F_k|}$ for $f \in E$. It is not hard to see that under these conditions a version of lemma 4.20 is valid, assuming that the W_{n-1}-boundary of W_n is small, in a manner depending on k, for all n such that $|W_n| > 2^{(h-\varepsilon/2)|F_k|}$, and this is just what quasi-incompressibility means. We omit the details.

5. Open questions

In conclusion we would like mention several problems which remain unresolved.

The method of proof of the fact that $T_* \geq h$ made it necessary that in the definition of R_k we require that a repetition of the pattern $x(F_k)$ be "counted" only if it is disjoint from the original pattern. This definition affects also the definition of $U_{k,n}$ and the results about U_*. In the case of $G = \mathbb{Z}^d$ studied in [8] no such restriction was necessary.

Question. *If we define*

$$R_k^{(F,W)}(x) = \inf\{n \,:\, \textit{there exists } 1_G \neq f \in W_n \textit{ s.t } fx \in x(F_k)\}$$

and $T_k = |W_{R_k}|$, *when is it true that* $\frac{1}{|F_k|}\log T_k \geq h$ *?*

For the upper bounds $T^* \leq h$ and $U^* \leq h$ to hold, we saw that it is necessary to require some special properties of the window sets, eg that they be (quasi) filling. We have seen that such sequences exist in many cases, but we have no information about the case of non-locally-finite torsion groups:

Question. *Do all groups possess (quasi) filling/incompressible sequences? What classes of groups have (quasi) filling/incompressible Følner sequences?*

Another issue is, under what conditions, milder than those given, does the bound $U_* \geq h$ hold? For one thing, the dependence on an interpolation sequence would seem to be an artifact of the proof, and perhaps can be removed. It is also not impossible that the requirement that $\{W_n\}$ be incompressible might be weakened to being filling, or perhaps even completely done away with; the fact that $T_* \geq h$ for very general window sequences, and that we have not yet found a counterexample of the type in section 2.7 for the lower bound $U_* \geq h$, gives some hope that this might be true.

Question. *Can weaker conditions on the window sequence be found which ensure the bound $U_* \geq h$?*

References

[1] H. Bass. The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc. (3)*, 25:603–614, 1972.

[2] Frederick P. Greenleaf. *Invariant means on topological groups and their applications*. Van Nostrand Reinhold Co., New York-Toronto, Ont.-London, 1969.

[3] Mikhail Gromov. Groups of polynomial growth and expanding maps. *Publ. Math. IHÉS*, 53:53–73, 1981.

[4] P. Hall and C. R. Kulatilaka. A property of locally finite groups. *J. London Math. Soc.*, 39:235–239, 1964.

[5] Elon Lindenstrauss. Pointwise theorems for amenable groups. *Inventiones Mathematicae*, 146:259–295, June 2001.

[6] Donald Ornstein and Benjamin Weiss. The Shannon-McMillan-Breiman theorem for a class of amenable groups. *Israel Journal of Mathematics*, 44(1):53–61, 1983.

[7] Donald Ornstein and Benjamin Weiss. Entropy and data compression schemes. *IEEE Transactions on Information Theory*, 39(1):78–83, January 1993.

[8] Donald Ornstein and Benjamin Weiss. Entropy and recurrence rates for stationary random fields. *IEEE Transactions on Information Theory*, 48(6):1694–1697, June 2002.

[9] Paul Shields. *The ergodic theory of discrete sample paths*, volume 13 of *Graduate Studies in Mathematics, ISSN 1065-7339*. American Mathematical Society, 1996.

[10] Benjamin Weiss. Monotileable amenable groups. *Amer. Math. Soc. Transl*, 202:257–262, 2001.

[11] Benjamin Weiss. Actions of amenable groups. In *Topics in dynamics and ergodic theory*, volume 310 of *London mathematical society Lecture Notes Series*. Cambridge university press, 2003.

[12] A. Wyner and J. Ziv. Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression. *IEEE Transactions on Information Theory*, 35:1250–1258, November 1989.

E-mail address: `mhochman@math.huji.ac.il`

Current address: Einstein Institute of Mathematics, Edmond J. Safra Campus, Givat Ram, The Hebrew University of Jerusalem Jerusalem, 91904, Israel